2017

I. Designing Brighter Fluorophores: A Computational And Spectroscopic Approach To Predicting Photophysical Properties Of Hydrazone-Based Dyes Ii. Developing Spectroscopic Methods To Better Understand The Cofactors Of Metalloproteins

Morgan Cousins
University of Vermont

I.
DESIGNING BRIGHTER FLUOROPHORES: A COMPUTATIONAL AND SPECTROSCOPIC APPROACH TO PREDICTING PHOTOPHYSICAL PROPERTIES OF HYDRAZONE-BASED DYES
II.
DEVELOPING SPECTROSCOPIC METHODS TO BETTER UNDERSTAND THE COFACTORS OF METALLOPROTEINS

A Dissertation Presented

by

Morgan E. Cousins

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Chemistry

October, 2017

Defense Date: August 7, 2017
Dissertation Examination Committee:

Matthew D. Liptak, Ph.D., Advisor
Adrian Del Maestro, Ph.D., Chairperson
Rory Waterman, Ph.D.
Matthias Brewer, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

Luminogens are molecules that emit light upon exposure to high-energy light, and fluorophores are one class of luminogens. Applications of fluorophores range from microviscosity sensors to light emitting diodes (LEDs), as well as biosensors, just to name a few. Many of these applications require the fluorophore to be in the aggregate or solid state. Some fluorophores become highly emissive in the aggregate state; these fluorophores are aggregation-induced emission (AIE) luminogens. Currently, very few quantum mechanical mechanisms have been proposed to describe the unique AIE behavior of luminogens.

Boron difluorohydrazone (BODIHY) dyes are a new type of AIE fluorophore. The bright emission is from the $S_{>1}$ excited state ("anomalous" emission) contrary to Kasha's Rule. Thus, the mechanism Suppression of Kasha's Rule (SOKR) was proposed to be responsible for the family of BODIHY dyes. We hypothesize that the SOKR mechanism can explain AIE as well as the anomalous emission of other fluorophores. New BODIHY derivatives (*para*-CO_2H BODIHY, aluminum difluorohydrazone (ALDIHY), and *para*-nitro ALDIHY) were predicted to be bright anomalous fluorophores through density functional theory (DFT) and time-dependent DFT (TDDFT) investigations. In addition, a series of anomalous fluorophores were investigated to determine if their photophysical properties could be explained by the SOKR mechanism (azulene, 1,6-diphenyl-1,3,5-hexatriene, and zinc tetraphenylporphyrin). Finally, several triazolopyridinium and triazoloquinolinium dyes were computationally investigated by DFT and TDDFT calculations, and an accurate computational model for the large Stokes shifts of these dyes was developed. In conclusion, a better understanding of the photophysical properties through DFT and TDDFT modeling and spectroscopic investigation of hydrazone-based fluorophores has been achieved.

In addition, the metal active sites and cofactors of metalloproteins were probed by optical spectroscopy, nuclear magnetic resonance (NMR) spectroscopy, and DFT modeling. In conjunction, these techniques can be used to elucidate the electronic structure responsible for the unique function of these metalloproteins. Specifically, a novel iron-sulfur cluster of a metalloprotein that may be involved in endospore formation of *Clostridium difficile*, CotA, was characterized by magnetic circular dichroism (MCD) spectroscopy. We propose that CotA contains a high-spin [4Fe-4S] cluster and a Rieske [2Fe-2S] cluster. It appears that the multimerization of the protein is related to the cluster conversion at the interface of monomeric subunits where two [2Fe-2S] clusters combine to form the [4Fe-4S] cluster. In addition, a putative cobalamin acquisition protein from *Phaeodactylum tricornutum*, CBA1, was not expressed at sufficient concentrations in *Escherichia coli* for spectroscopic investigation. Finally, a new technique was developed using cobalt-59 NMR spectroscopy to better understand the nucleophilic character of cobalt tetrapyrroles, such as cobalamin (vitamin B12), as biological cofactors as well as synthetic catalysts. New insight into the electronic structure provides valuable information related to the mechanism of these metalloproteins.

CITATIONS

Material from this dissertation chapter has been published in the following form:

Qian, H., Cousins, M. E., Horak, E. H., Wakefield, A., Liptak, M. D., Aprahamian, I.. (2017). Suppression of Kasha's rule as a mechanism for fluorescent molecular rotors and aggregation-induced emission. *Nat. Chem.*, **9**, 83-97.

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

Abs .....UV/vis absorption

ACQ .....aggregation caused quenching

Ado.....5’-deoxy-5’-adenosine

AdoCbl .....Adenosylcobalamin

AIE .....aggregation induced emission

Az.....azulene

B3LYP .....Becke, three-parameter, Lee-Yang-Parr

bisImCbl.....bisimidazolylcobalamin

BLYP .....Becke- Lee-Yang-Parr

BODIHY .....boron difluorohydrazone

BSA.....bovine serum albumin

CBA1 .....cobalamin acquisition protein 1

Cbl.....Cobalamin

CD .....circular dichroism

CD630 .....*Clostridium difficile* pathogenic strain 630

CIE .....Commission Internationale de L'éclairage

CNCbl .....cyanocobalamin, vitamin B12

CPD.....cysteine protease domain

CT .....charge transfer

DCM .....dichloromethane

DCVJ.....9-(dicyanovinyl) julolidine

DET.....Dexter energy transfer

DFT.....density functional theory

diCNCbl.....dicyanocobalamin

DMB 5,6-dimethylbenzimidazole

DMSO.....dimethylsulfoxide

DPH.....1,6-diphenyl-1,3,5-hexatriene

DSS .....sodium 2,2-dimethyl-2-silapentane-5-sulfonate

DTT.....dithiothreitol

ED .....electron donating

EDTA.....ethylenediaminetetraacetic acid

EG .....ethylene glycol

EtOH .....ethanol

EW .....electron withdrawing

E_{XC} .....exchange-correlation energy

FID .....free induction decay

FMR .....fluorescent molecular rotor

f_{osc}.....oscillator strength

FRET.....Förster resonance energy transfer

GGA.....generalized gradient approximation

Gly.....glycerol

H_2OCbl.....aquocobalamin, vitamin B12a

Hcy.....Homocysteine

HOCbl .....hydroxocobalamin

HOMO .....highest occupied molecular orbital

IC.....internal conversion

Im .....imidazolate

ImCbl .....imidazolylcobalamin

ImH .....imidazole

IPTG.....isopropyl β-D-1-thiogalactopyranoside

ISC .....intersystem crossing

ISC .....iron-sulfur cluster

IUPAC.....International Union of Pure and Applied Chemistry

LB .....lysogeny broth

LE.....locally excited

LED.....light emitting diode

LFS.....ligand field splitting

LMCT .....ligand-to-metal charge transfer

LUMO.....lowest unoccupied molecular orbital

MCD .....magnetic circular dichroism

MeCbl .....Methylcobalamin

Met .....Methionine

MetTHF.....methyltetrahydrofolate

MMCM .....methylmalonyl-CoA mutase

MO .....molecular orbital

NMR .....nuclear magnetic resonance

OLED.....organic light emitting diode

PBE .....Perdew-Burke-Erzerhof

PES.....potential energy surface

PET .....photoinduced electron transfer

PS-EG .....pre-stained ethylene glycol

RIM.....restriction of intramolecular motion

RIR.....restriction of intramolecular rotation

RIV.....restriction of intramolecular vibration

$S_{>1}$ .....singlet excited state higher in energy than S_1

S_0.....singlet ground state

S_1.....lowest energy singlet excited state

SDS .....sodium dodecyl sulfate

SDS-PAGE .....sodium dodecyl sulfate polyacrylamide gel electrophoresis

SOC.....spin-orbit coupling

SOKR.....Suppression of Kasha's Rule

T_1.....lowest energy triplet excited state

TB .....terrific broth

TDDFT.....time-dependent density functional theory

TEM .....transmission electron microscopy

TEV.....tobacco etch virus

THF .....tetrahydrofuran

TICT.....twisted intramolecular charge transfer

TMS .....tetramethylsilane

TOP .....[1,2,3]triazolo[1,5-a]pyridinium

TOQ ..... [1,2,3]triazolo[1,5-a]quinolinium

TPE .....tetraphenylethene

TPP.....5,10,15,20-tetraphenylporphin

TPSS .....Tao−Perdew−Staroverov−Scuseria

UVM .....University of Vermont

Zn TPP .....zinc 5,10,15,20-Tetraphenyl-21*H*,23*H*-porphine

LIST OF TABLES

Table **Page**

LIST OF FIGURES

LIST OF SCHEMES

Scheme **Page**

CHAPTER 1: INTRODUCTION

SECTION 1.1. THE ELECTRONIC STRUCTURE OF HYDRAZONE-BASED FLUOROPHORES AND METALLOPROTEINS GREATLY INFLUENCES THEIR FUNCTION.

The photophysical properties of a fluorescent molecule are rooted in the electronic structure of the fluorophore. The catalytic function of metalloproteins arises from the electronic structure of the active site. The electronic structure of a fluorophore or metalloprotein can be elucidated either spectroscopically or computationally. The use of these techniques in tandem provides a more complete description of the electronic structure.

Several spectroscopic techniques will be used to directly probe electronic transitions between the ground and excited states of hydrazone-based fluorophores in this dissertation: fluorescence spectroscopy, fluorescence lifetime spectroscopy, UV/vis absorbance (Abs) spectroscopy, and circular dichroism (CD) spectroscopy. To understand the electronic structure of the metalloproteins' cofactors, Abs, CD, magnetic circular dichroism (MCD) and nuclear magnetic resonance (NMR) spectroscopy will be used to describe their electronic structure. In addition, the electronic structure of these hydrazone-based dyes as well as metalloproteins will be predicted using density functional theory (DFT) and time-dependent DFT.

SECTION 1.2. JABLONSKI DIAGRAM DESCRIBES LUMINESCENCE

Typically, a molecule will absorb high energy incident light then, through a series on non-radiative relaxation pathways, reach the lowest energy excited state of a given multiplicity (**Fig. 1.1**). This behavior follows Kasha's Rule which describes fluorescence and phosphorescence of most luminogens.[1] Fluorescence is the radiative relaxation from a singlet excited state and phosphorescence is radiative relaxation from a triplet excited state. The triplet excited state is accessible through intersystem crossing (ISC) leading to phosphorescence. ISC can be enhanced by the presence of heavy atoms.[2] This dissertation will primarily discuss fluorophores.

Following excitation into a high-lying excited state, some fluorophores will relax radiatively from an excited state other than S_1.[3] Radiative relaxation can occur from a higher energy excited state as well as high-lying vibrational states (**Fig. 1.1**). This is "anomalous" emission and instances of this have been documented since the 1950s with azulene.[4] The higher energy emission from the $S_{>1}$ excited states is attractive for many types of applications.

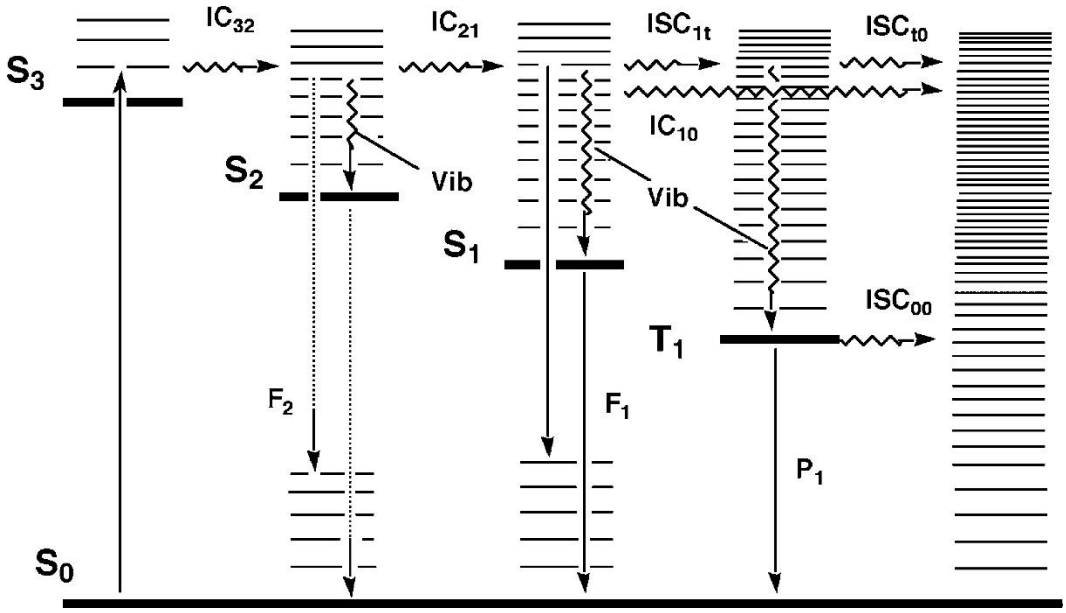


Figure 1.1. A Jablonski diagram illustrating the non-radiative and radiative processes following excitation from the ground state. The abbreviations are as follows: S_0, ground state; S_1, S_2, and S_3, singlet exited states; T_1, triplet excited state; IC, internal conversion; Vib, vibrational relaxation; ISC, intersystem crossing; F_1, fluorescence; F_2, anomalous emission; P_1, phosphorescence. Taken from reference 3.

SECTION 1.3. APPLICATIONS OF FLUOROPHORES

There are a wide variety of uses for luminogens attracting interest from the fields outside of chemistry such as biology, medicine, and materials science (**Fig. 1.2**). The fluorophores can be tailored to suit a specific function through structural modification. Many applications use thin-film or aggregate-state fluorophores.[5]

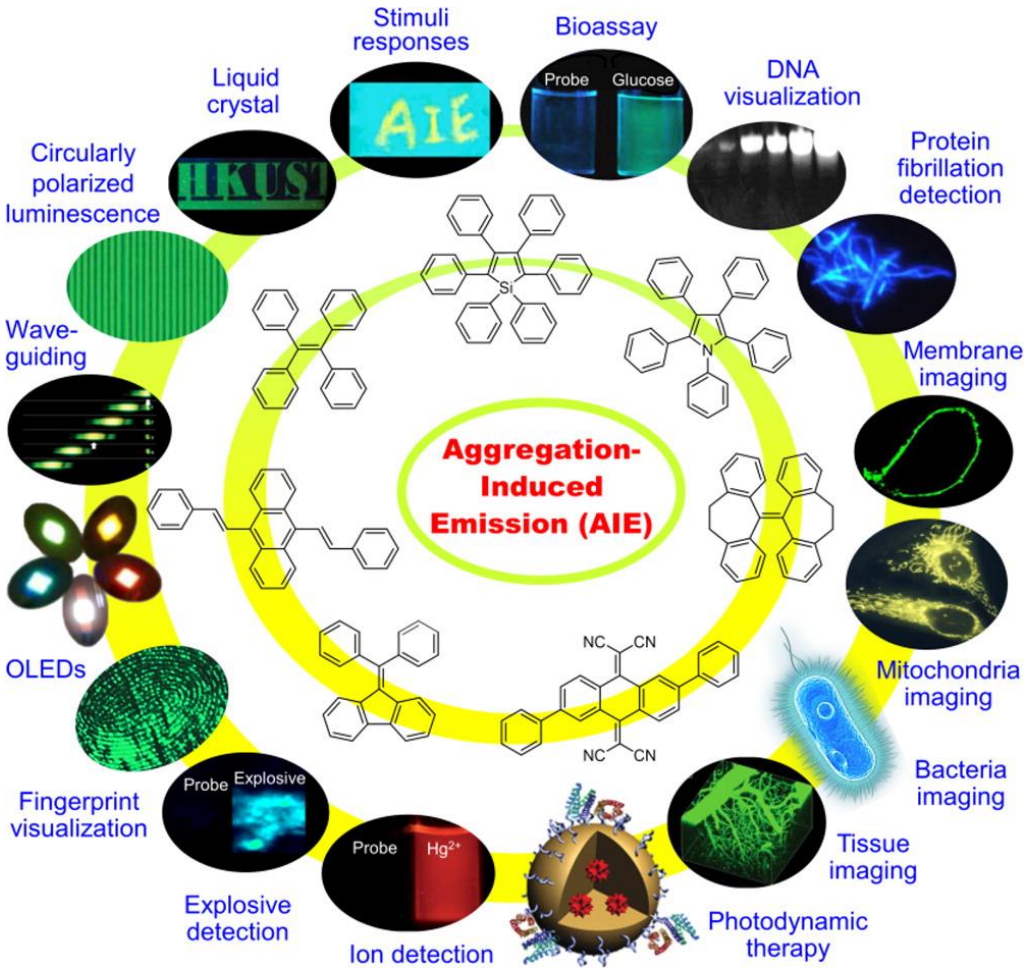


Figure 1.2. Examples of the many and varied applications of luminogens that experience emission enhancement upon aggregation. This figure has used and adapted images from the following references.[6] Figure taken from reference 5.

For the biological and detection applications, aqueous solvent would be preferred. The addition of hydrophilic groups can increase water solubility of these luminogens.[7] However, many fluorophores are organic and hydrophobic due to large aromatic moieties. The precipitation of these fluorophores often results in decreased emission. However, the fluorophores that experience emission enhancement in the aggregate state have opened the door for new applications of these molecules especially in the field of biology. One advancement of note is the biological imaging capabilities that have greatly improved due to the discovery of fluorophores that are brighter in the aggregate state (**Fig. 1.1**).

For example, tetraphenylethene (TPE) has been investigated for its potential use as a biological probe.[8] This fluorophore had a large Stokes shift (almost 9000 cm^{-1}). This fluorophore was not emissive in water but was highly emissive in non-polar solvents as well as solutions with bovine serum albumin (BSA) (**Fig. 1.3**). The dye selectively complexed with native BSA in a fold-sensitive manner.[8] The turn-on probe can detect BSA at micromolar concentration. This fluorophore could be used to detect hydrophobic regions on proteins as it was proposed to bind in a hydrophobic pocket in BSA based on SDS-denaturing studies. This fluorophore was used in later studies to understand enzymatic hydrolysis as sugar chains were added to the phenyl at the para-position.[9]

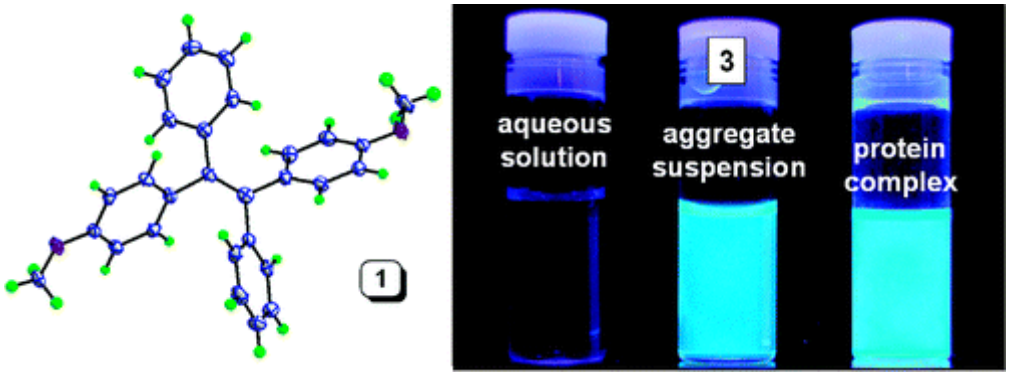


Figure 1.3. Enhanced emission was observed for TPE in the aggregate state and protein-bound. (**Left**) The crystal structure of methoxyphenyl derivative of TPE. (**Right**) Vials containing (3-sulfonatopropoxyl) phenyl derivative of TPE: dissolved in water, dissolved in MeCN: water (99: 1) mixture, and complexed with BSA in phosphate buffer at pH 7. Figure taken from reference 8.

Attractive attributes for a fluorophore for industrial and research applications include high quantum yield, large Stokes shift, and tunable emission. Efficient fluorophores have high quantum yields which are directly correlated to brightness. A large Stokes shift prevents lowering the efficacy of the dye arising from reabsorption of the emission.[10] The color and efficacy of emission can be altered synthetically by the addition of new functional groups or aromatic substitutions.

SECTION 1.4. INFLUENCE OF AGGREGATION ON EMISSION

Most luminogens are used as a thin-film or the aggregate state for industrial applications. The formation of aggregates can drastically influence the emission of the luminogen and is typically an irreversible process.[11] Aggregation can lead to quenching (aggregation caused quenching, ACQ) or increase (aggregation induced emission, AIE) of emission. Some luminogens are not effected by aggregation; however, most fluorophores will undergo ACQ (**Fig. 1.4**).[12]

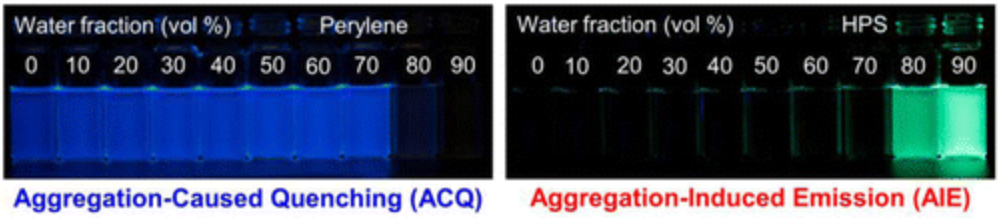


Figure 1.4. The quenching and enhancing effects of aggregation on two fluorophores. (**Left**) perylene (20 μM) in solutions of water and tetrahydrofuran (THF). Emission is greatly diminished upon aggregation (80% water, 20% THF) due to ACQ. (**Right**) hexaphenylsilole (20 μM) in solutions of water and THF. Upon aggregation, the emission of hexaphenylsilole is greatly enhanced due to AIE effects. Figure taken from reference 5.

Deleterious intramolecular interactions lead to ACQ. ACQ can occur due to high concentrations, solid-state, or solvent-caused aggregation. Intramolecular interactions that result in ACQ include excimer formation and π-stacking. ACQ is common for aromatic fluorophores where π-stacking occurs between molecules.[5] These intramolecular interactions can be interrupted if bulky groups are added to the ACQ fluorophore to interrupt the packing interaction. The mechanism of ACQ is well-understood; these

intramolecular interactions lead to non-radiative relaxation or decay pathways to the ground state.[7]

AIE luminogens were first reported in 2001.[13] As this is a fairly new discovery, the quantum mechanical mechanism of AIE is far less understood as compared to ACQ. The restriction of intramolecular motion (RIM) phenomenon is often observed for emission enhancement; intramolecular motions include rotation and vibrations.[6a] Because aggregation is a requirement for so many industrial applications of luminogens, AIE fluorophores are suited for more applications.

There are two common motifs that arise from aggregation of fluorophores: J-aggregates and H-aggregates. The formation of J-aggregates leads to a lower-energy (red-shift) in emission intensity.[14] The J-aggregates do not completely overlap but instead assemble in a staggered head-to-tail orientation; this enhances the rate of fluorescence leading to emission enhacement.[15] The formation of J-aggregates leads to AIE while H-aggregates cause ACQ. The H-aggregates will assemble with complete overlap in a stacking fashion.[15] The H-aggregation affects the emission spectra by reducing the radiative decay rate and shifting the emission to a higher-energy.[15]

SECTION 1.5. DISTRIBUTION OF METALS IN BIOLOGY

Metal binding is vital for function for many enzymes. A variety of elements have been shown to be bound my enzymes (**Fig. 1.5**). The metal cofactor is often bound to the protein by an evolutionarily conserved motif as determined by proteomics.[16] The enzymatic role of proteins containing metals can be found everywhere in nature. This dissertation will exclusively discuss transition metal binding by proteins (specifically iron and cobalt).

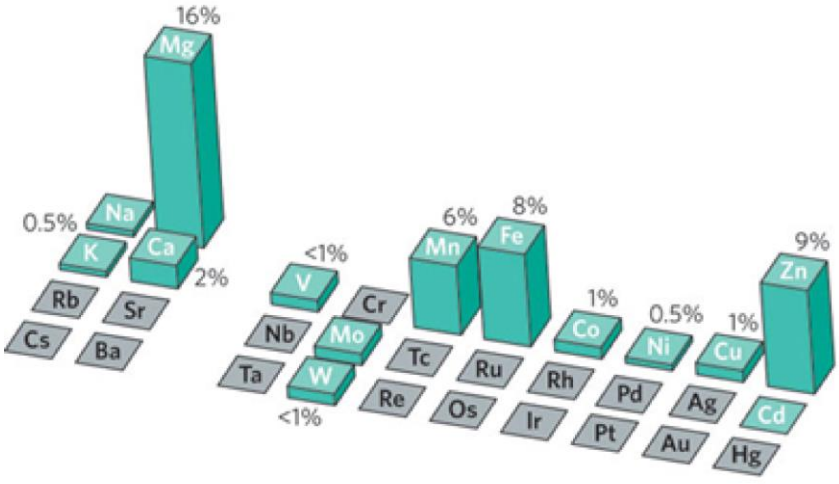


Figure 1.5. The natural occurrence of elements used by enzymes as cofactors. The height of the column represents the percentage of all enzymes that use that element. Taken from reference 3.

Amino acids bind to the metal ion(s) in the active site of the protein through N, O, or S atoms (**Fig. 1.6**). The amino acids highlighted in red in **Figure 1.6** are the most likely ligands for the iron atoms in iron-sulfur clusters as well as the cobalt atom in cobalamin; histidine (HIS) binds to the cobalt through the nitrogen of the imidazole group following

deprotonation while cysteines (CYS) bind to the iron atoms through the thiolate groups.[17] The catalytic activity of the metal active site is controlled by the primary sphere (bound amino acids) and the secondary sphere (amino acids in the binding site that interact with metal, cofactor, or first-sphere ligand through non-binding interactions).

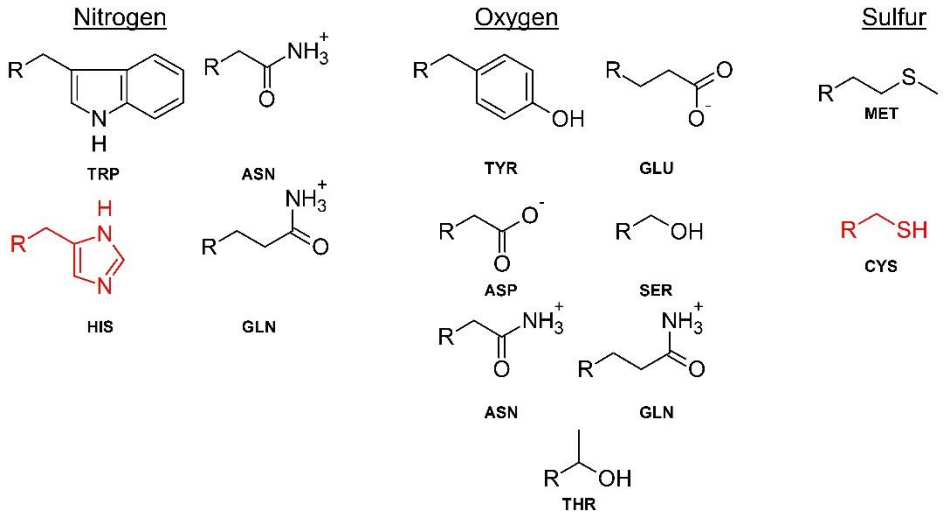


Figure 1.6. Amino acids that bind to metals at pH 7.0 organized by binding atom. The R group represents the protein amide backbone. The abbreviations represent the following naturally occurring amino acids: TRP, tryptophan; ASN, asparagine; HIS, histidine (pKa 6.0); GLN, glutamine (pKa 6.04); TYR, tyrosine (pKa 10.5); GLU, glutamate (pKa 4.1); ASP, aspartate (pKa 3.9); SER, serine; THR, threonine; MET, methionine; and CYS, cysteine (pKa 8.4)). ASN and GLN can bind through both the O and N atom but preferentially bind to metals through the oxygen atom.

The diversity of functionality for metalloproteins spans all six Enzyme Commission (EC) classes.[18] Iron is the most prevalent transition metal and biology and most iron-containing enzymes are oxidoreductases (**Fig. 1.7**). Iron is used in a variety of cofactors

such as divalent centers, iron sulfur cluster, and heme all of which play an integral role in redox chemistry.[19] Not all metals are used in biochemical reactions. For example, zinc is required for the proper folding of structural motifs such as zinc-fingers.[18]

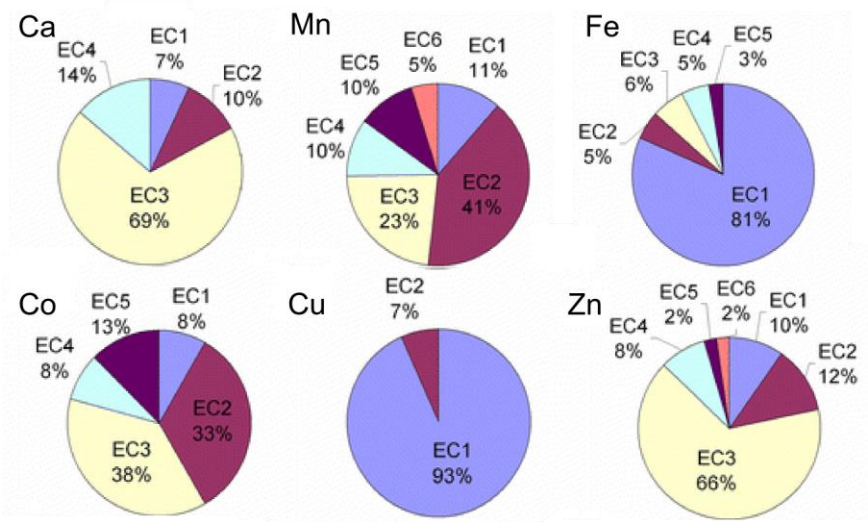


Figure 1.7. These pie charts show the distribution of the functions of enzymes that bind a row 3 metal ion. The EC numbers represent the following: EC1, oxidoreductases (blue); EC2 transferases (magenta); EC3, hydrolases (yellow); EC4, lyases (green); EC5, isomerases (purple); and EC6, ligases (pink). Adapted from reference 4.

The function of these enzymes is heavily regulated by the metal active site. To understand the catalysis of metalloenzymes, the electronic structure of the metal site must be well understood. The structure of the entire metalloprotein, while useful, cannot provide information to understand the catalysis. In the following chapters, the cobalt-containing cofactor cobalamin as well as iron sulfur clusters will be characterized spectroscopically. Their electronic structures will be investigated through many spectroscopic techniques: Abs, CD, MCD, and NMR.

SECTION 1.6. DIVERSITY OF IRON-SULFUR CLUSTERS

Iron-sulfur clusters (ISCs) are diverse in structure and function. ISCs are the oldest bioinorganic cofactor as proposed by the "iron-sulfur world theory".[20] The diversity in the cluster type is reflective of the multitude of functions of ISCs. Some functions of ISCs include electron transfer, coupled electron/proton transfer, storage for iron or cluster, as well as substrate binding and activation.[21] As free iron would participate in damaging Fenton chemistry to product reactive oxygen species, the metal is likely delivered to the protein by a chaperone protein. In the bacteria *Azotobacter vinelandii*, two systems (assembly and scaffold) are involved in ISC assembly.[22] However, *in vitro* reconstitution of the ISC can be achieved upon the addition of sulfide and iron during expression or purification.[23]

In ISC, cysteine is the most common amino acid to bind the iron atom(s) while other amino acids (histidine, aspartate, arginine, and serine) can bind to the iron atoms in Rieske clusters.[16] The rhombic [2Fe-2S] and cubane [4Fe-4S] clusters are the most prevalent ISCs.[24] There are many variations on the FeS cluster motif and each structure has many accessible oxidation states and spin states (**Fig. 1.8**). A paramagnetic oxidation state is more common which allows for characterization through a multitude of spectroscopic techniques.

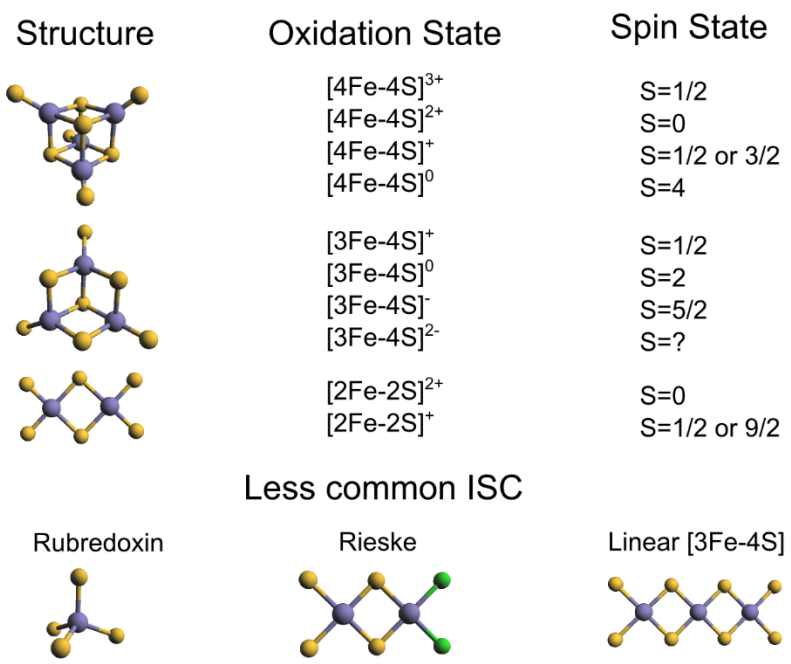


Figure 1.8. Some iron-sulfur clusters with their potential oxidations states and spin states. The structure of [4Fe-4S] cluster is cubane, the structure of the [3Fe-4S] cluster is pseudo cubane, and the [2Fe-2S] is a linear cluster. Silver spheres represent iron (accessible oxidation states include Fe^{2+}, $Fe^{2.5+}$, and Fe^{3+}). Golden spheres represent sulfur (cysteine or sulfide). Green spheres represent non-sulfur ligands (such as nitrogen from histidine) in a Rieske cluster.

Characterizing iron-sulfur clusters with spectroscopic techniques. Many spectroscopic techniques have been used to identify and characterize FeS clusters of a metalloprotein of interest. The variety of useful techniques stems from the inherent magnetism and electron-rich structure of the many ISC. Two methodologies used to probe metalloproteins containing ISC successfully are Mössbauer spectroscopy and electron paramagnetic resonance (EPR) spectroscopy.[25]

EPR spectroscopy can provide information pertaining to the structure and function of many types of ISC proteins and was the first technique used to investigate ISC proteins.[26] This technique can be used to provide the following pieces of information pertaining to the ISC present in a metalloprotein: redox states, electronic structure, as well as proximity of many clusters (if multiple clusters are present).[26] EPR spectroscopy can characterize the spin state of the cluster; in the typical perpendicular configuration, a spectrum without resonances would indicate an integer-spin species.[27] Mössbauer spectroscopy is inherently specific to the FeS cluster ISC proteins because the technique relies of the Fe-57 isotope (a naturally abundant isotope, 2.1%).[24] This technique provides information on the oxidation and spin state of the iron atom(s) of the cluster.[27]

Multiple techniques are often required to be used in tandem to characterize an ISC protein. Other methods used to study FeS proteins include electron–nuclear double resonance (ENDOR) spectroscopy, electron spin echo envelope modulation (ESEEM) spectroscopy, X-ray absorption spectroscopy, CD spectroscopy, and MCD spectroscopy. The simplest experiment to collect information pertaining to the presence of the ISC is Abs spectroscopy as CT features (sulfur-to-iron) are easily observed due to the intensity of the features ($1 \leq \varepsilon \leq 100$ M^{-1} cm^{-1}).[27] This dissertation will describe the characterization of an ISC by Abs and MCD spectroscopy.

SECTION 1.7. FUNCTIONS OF COBALAMIN CONTAINING ENZYMES

Cobalamin is a biologically relevant cobalt tetrapyrrole. Cobalt is almost exclusively found in metalloproteins as cobalamin.[19] The cobalamin family have a conserved cobalt-bound tetrapyrrole structure (**Fig. 1.9**). There are three oxidation states for Cbl: Co(I), Co(II), and Co(III). The Co(I) cobalamins are supernucleophilic.[28] The accessibility of these oxidation states allow the enzyme-bound cobalamins to complete a variety of different chemical reactions.

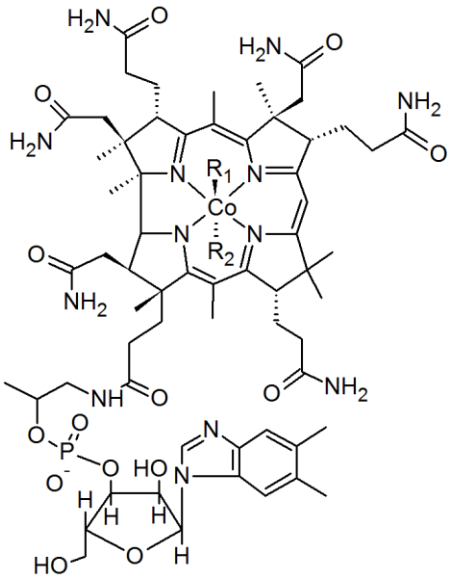


Figure 1.9. General structure of cobalamin family. Cobalamin, also known as vitamin B12, is a cobalt containing corrin ring with variable upper β-axial ligand, "R_1". The biologically relevant forms of cobalamin are methylcobalamin (R_1= methyl), adenosylcobalamin (R_1= 5'-deoxyadenosyl), cyanocobalamin (R_1= CN), and aquocobalamin or hydroxocobalamin (R_1= OH_2 and R= OH). The α-axial binding site (R_2) can be the nucleotide base, 5,6-dimethylbenzimidazole (DMB), in the "base-on" conformation or another ligand, such as histidine or water, in the "base-off" conformation.

Vital enzymatic processes utilize cobalamin. Archaea and bacteria at the bottom of the food chain produce this micronutrient while other organisms must obtain cobalamin through their diet or environment.[29] Organisms from the Bacteria, Archaea, and Eukaryota domains require cobalamin for life-dependent pathways such as methionine and L-methylmalonyl-CoA synthesis. Cobalamin-dependent enzymes perform diverse functions and include isomerases, methyltransferases, and reductive halogenases.[30] The vital biochemical roles of AdoCbl and MeCbl lie in the stability Co—C bond of these B12 coenzymes. Cobalamin-dependent enzymes include methionine synthase and methylmalonyl-CoA mutase (MMCM).

Methylcobalamin is required for one-carbon metabolism. Methylcobalamin (MeCbl) is utilized as a cofactor for enzymes responsible for carbon skeletal rearrangements.[17] Enzymes that require MeCbl include methyltransferases such as methionine synthase.[31] Methionine synthase requires cobalamin as an intermediate methyl group carrier prior to the heterolytic cleavage of the Co(III)—C bond.[17] This enzyme is responsible for producing methionine (Met) as well as tetrahydrofolate (THF) from homocysteine (Hcy) and 5-methyltetrahydrofolate (MeTHF) (**Fig. 1.10**).

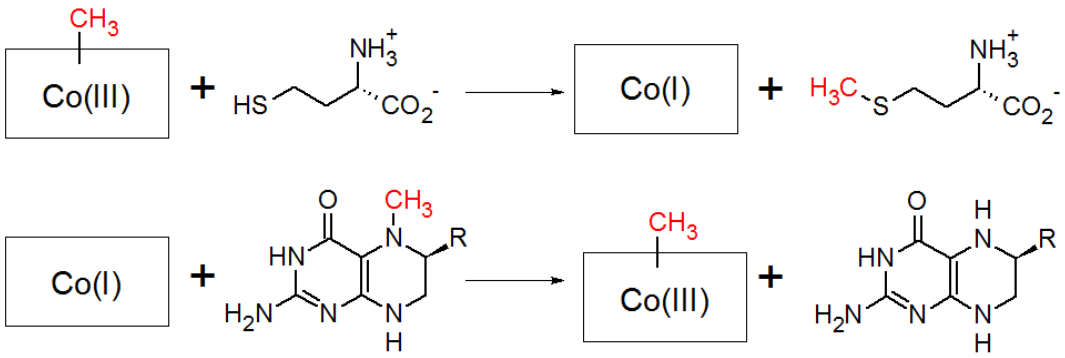


Figure 1.10. The two consecutive S_N2 reactions of the cobalamin-dependent methionine synthase pathway. Methylcobalamin serves as the methyl carrier where the methyl is removed by heterolytic cleavage producing the cob(I)alamin species. Cbl receives a methyl group from MeTHF to produce THF and restart the catalytic cycle.

The supernucleophilic character of cob(I)alamin allows for the mechanistically challenging reaction described in **Figure 1.10** to occur. The transition state must be comprised of a linear bond between the cobalt center, methyl group and leaving group and the cleavage reactions are concerted for this S_N2 reaction.[32] Then cob(I)alamin is shuttled to a second site in the protein where the methyl group is transferred to the sulfur group of deprotonated Hcy.[33] The catalytic cycle restarts after a methyl group donation to cob(I)alamin from MeTHF. The oxidation of Co(I)Cbl will inactivate the cofactor; a reductive methylation by flavodoxin and S-adenosyl-L-methionine is required to reactivate the enzyme.[32]

Two derivatives of the methionine synthase enzyme support this single-carbon metabolism pathway: cobalamin-dependent (MetH) or cobalamin-independent (MetE) methionine synthase.[17] The cobalamin-dependent pathway is more efficient; the turnover number for MetH is 1500 min^{-1} but only 12.3 min^{-1} for MetE.[34] Organisms without a

cobalamin-uptake pathway encode the *metE* gene.[35] Organisms that synthesize or readily obtain cobalamin from their environment usually express the cobalamin-dependent enzyme.[35] Bacteria, such as *Escherichia coli*, will express both MetE and MetH.[35] Organisms possessing both forms of the enzyme often favor the cobalamin-dependent enzyme but can be flexible in their demand for cobalamin.[29b]

Adenosylcobalamin is required for methylmalonyl-CoA mutase activity. Adenosylcobalamin (AdoCbl) is the second biologically relevant cobalamin derivative required for enzymes such as MMCM, diol dehydrase, and L-leucine 2,3-aminomutase.[31] Also known as coenzyme B12, this cofactor is homolytically cleaved to form a carbon-based radical and a paramagnetic Co(II) center in these isomerizations (**Fig. 1.11**).

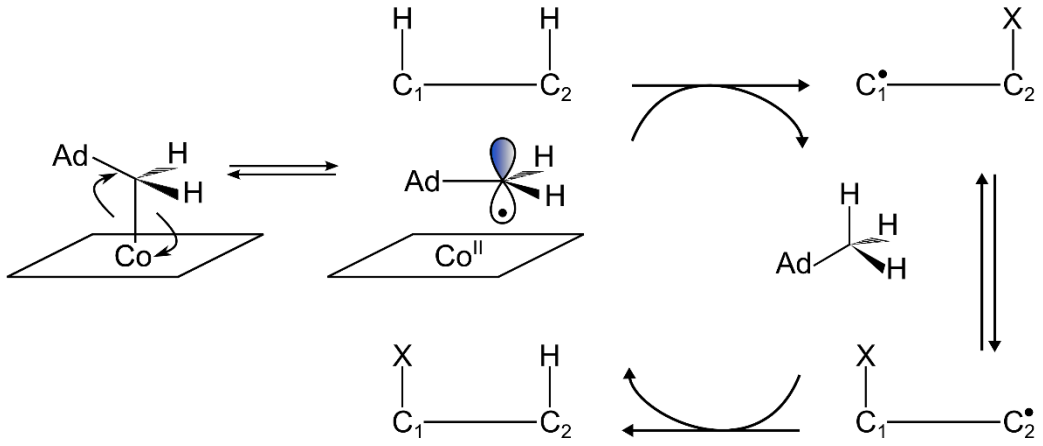


Figure 1.11. Generalized mechanism of AdoCbl-dependent enzymatic isomerizations. homolytic cleavage of Co(III) —C bond produces 5'-deoxy-5'-adenosine (Ado) radical and cob(II)alamin. Ado transfers radical to substrate which undergoes a radical rearrangement. Radical is quenched with back-transfer of H-atom to Ado and reformation of AdoCbl. Adapted from reference 4.

The covalent bond ligating the Ado moiety to the Co(III) center is relatively weak: the bond dissociation energy of the Co—C bond is 30 kcal mol^{-1}.[36] This radical formation is often involved in isomerization reactions that require the cleavage of a carbon-carbon, carbon-oxygen, or carbon-nitrogen bond.[37] Methylmalonyl-CoA mutase relies on AdoCbl as a cofactor for fatty acid metabolism (**Figure 1.12**).

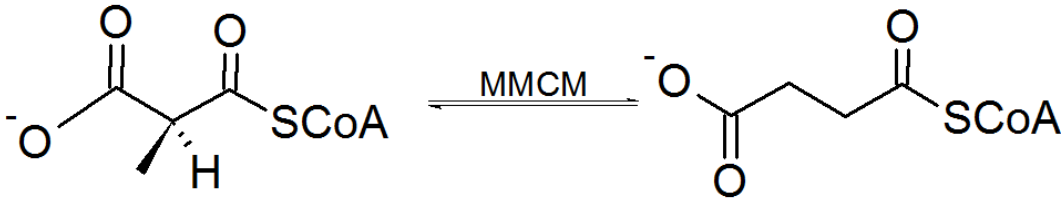


Figure 1.12. Enzymatic mechanism of methylmalonyl-CoA mutase. L-methylmalonyl-CoA is converted to Succinyl-CoA through a radical rearrangement.

SECTION 1.8. EXPERIMENTAL APPROACH TO CHARACTERIZE THE ELECTRONIC STRUCTURE OF HYDRAZONE-BASED FLUOROPHORES AND BIOINORGANIC COFACTORS

In chapters 2, anomalous emission of boron difluorohydrazone (BODIHY) dyes will be detected using a suite of spectroscopic techniques: fluorescence spectroscopy, fluorescence lifetime spectroscopy, Abs spectroscopy, and CD spectroscopy. These techniques will provide more information pertaining to the ground and excited states of the fluorophore. The influence of solvent polarity and viscosity will be investigated to determine the photophysical mechanism(s) behind the unique emission. The BODIHY dyes, as well as several known fluorophores with anomalous emission, will be subjected to these investigations in chapters 2 and 4.

In combination with these techniques, the photophysical properties of these dyes will be modeled with density functional theory (DFT) and time-dependent DFT (TDDFT). A functional will be assessed on its ability to predict the Abs spectrum of a fluorophore; this will ensure the function is accurately modeling the electronic structure of the fluorophore. Derivatives of BODIHY will be computationally investigated to predict brightness, behavior as a molecular rotor, as well as the influence the para-substituent on the pendant phenyl can have on the emission wavelength in chapter 3. A large Stokes shift is an attractive attribute of a fluorophore as it prevents reabsorption. The triazolopyridinium and triazoloquinolinium dyes have large Stokes shifts (over 7000 cm^{-1}) which have been modeled by TDDFT in chapter 5.

In chapter 6, a combination of Abs, CD and MCD spectroscopy will be used to characterize an unknown iron-sulfur cluster bound by a metalloprotein in *Clostridium difficile*. MCD spectroscopy is sensitive to the cluster type and oxidation state while being insensitive to the protein structure beyond the active site. A library of published MCD spectra specific to ISC structures and oxidation states will be used to characterize the ISC of CotA.

In chapter 7, the expression in *Escherichia coli* and purification of an algal cobalamin acquisition protein, CBA1, will be discussed. A new technique to assess the nucleophilic character of cobalamins will be discussed in chapter 8. This methodology uses ^{59}Co NMR spectroscopy to better understand catalytic activity of cobalamin-dependent enzymes.

SECTION 1.9. CHAPTER 1 REFERENCES

1. Kasha, M., *Disc. Faraday Soc.* **1950,** *9*, 14-19.

2. McGlynn, S. P.; Daigre, J.; Smith, F. J., *J. Chem. Phys.* **1963,** *39*, 675.

3. Itoh, T., *Chem. Rev.* **2012,** *112*, 4541-4568.

4. Beer, M.; Longuet-Higgins, H. C., *J. Chem. Phys.* **1955,** *23*, 1390.

5. Mei, J.; Leung, N. L.; Kwok, R. T.; Lam, J. W.; Tang, B. Z., *Chem. Rev.* **2015,** *115*, 11718-11940.

6. (a) Mei, J.; Hong, Y.; Lam, J. W.; Qin, A.; Tang, Y.; Tang, B. Z., *Adv. Mater.* **2014,** *26*, 5429-5479; (b) Liu, Y.; Deng, C.; Tang, L.; Qin, A.; Hu, R.; Sun, J. Z.; Tang, B. Z., *J. Am. Chem. Soc.* **2011,** *133*, 660-663; (c) Hong, Y.; Chen, S.; Leung, C. W.; Lam, J. W.; Tang, B. Z., *Chem. Asian J.* **2013,** *8*, 1806-1812; (d) Hong, Y.; Meng, L.; Chen, S.; Leung, C. W.; Da, L. T.; Faisal, M.; Silva, D. A.; Liu, J.; Lam, J. W.; Huang, X.; Tang, B. Z., *J. Am. Chem. Soc.* **2012,** *134*, 1680-1689; (e) Zhang, C.; Jin, S.; Yang, K.; Xue, X.; Li, Z.; Jiang, Y.; Chen, W. Q.; Dai, L.; Zou, G.; Liang, X. J., *ACS Appl. Mater. Interfaces* **2014,** *6*, 8971-8975; (f) Zhao, N.; Li, M.; Yan, Y. L.; Lam, J. W. Y.; Zhang, Y. L.; Zhao, Y. S.; Wong, K. S.; Tang, B. Z., *J. Mater. Chem. C* **2013,** *1*, 4640-4646; (g) Zhao, E. G.; Hong, Y. N.; Chen, S. J.; Leung, C. W. T.; Chan, C. Y. K.; Kwok, R. T. K.; Lam, J. W. Y.; Tang, B. Z., *Advanced Healthcare Materials* **2014,** *3*, 88-96; (h) Ding, D.; Goh, C. C.; Feng, G. X.; Zhao, Z. J.; Liu, J.; Liu, R. R.; Tomczak, N.; Geng, J. L.; Tang, B. Z.; Ng, L. G.; Liu, B., *Adv. Mater.* **2013,** *25*, 6083-6088; (i) Yuan, Y. Y.; Feng, G. X.; Qin, W.; Tang, B. Z.; Liu, B., *Chem. Commun.* **2014,** *50*, 8757-8760; (j) Zhao, N.; Lam, J. W. Y.; Sung, H. H. Y.; Su, H. M.; Williams, I. D.; Wong, K. S.; Tang, B. Z., *Chem. Eur. J.* **2014,** *20*, 133-138; (k) Guo, Y. X.; Feng, X.; Han, T. Y.; Wang, S.; Lin, Z. G.; Dong, Y. P.; Wang, B., *J. Am. Chem. Soc.* **2014,** *136*, 15485-15488; (l) Li, Y.; Xu, L. R.; Su, B., *Chem. Commun.* **2012,** *48*, 4109-4111; (m) Zhao, D.; Fan, F.; Cheng, J.; Zhang, Y.; Wong, K. S.; Chigrinov, V. G.; Kwok, H. S.; Guo, L.; Tang, B. Z., *Adv. Opt. Mater.* **2015,** *3*; (n) Yuan, W. Z.; Tan, Y.; Gong, Y.; Lu, P.; Lam, J. W. Y.; Shen, X. Y.; Feng, C.; Sung, H. H. Y.; Lu, Y.; Williams, I. D., *Adv. Mater.* **2013,** *25*; (o) Liu, J.; Su, H.; Meng, L.; Yihua, Z.; Chunmei, D.; Ng, J. C. Y.; Lu, P.; Faisal, M.; Lam, J. W. Y.; Huang, X.; Wu, H.; Wong, K. S.; Tang, B. Z., *Chem. Sci.* **2012,** *3*, 2737-2747.

7. Hong, Y.; Lam, J. W.; Tang, B. Z., *Chem. Commun.* **2009**, 4332-4353.

8. Tong, H.; Hong, Y.; Dong, Y.; Haussler, M.; Li, Z.; Lam, J. W.; Dong, Y.; Sung, H. H.; Williams, I. D.; Tang, B. Z., *J. Phys. Chem. B* **2007,** *111*, 11817-11823.

9. Wang, J. X.; Chen, Q.; Bian, N.; Yang, F.; Sun, J.; Qi, A. D.; Yan, C. G.; Han, B. H., *Org. Biomol. Chem.* **2011,** *9*, 2219-2226.

10. Hong, Y. N., *Methods and Applications in Fluorescence* **2016,** *4*, 1-17.

11. Zhai, D.; Xu, W.; Zhang, L.; Chang, Y. T., *Chem. Soc. Rev.* **2014,** *43*, 2402-2411.

12. Ma, X. F.; Sun, R.; Cheng, J. H.; Liu, J. Y.; Gou, F.; Xiang, H. F.; Zhou, X. G., *J. Chem. Educ.* **2016,** *93*, 345-350.

13. Luo, J.; Xie, Z.; Lam, J. W.; Cheng, L.; Chen, H.; Qiu, C.; Kwok, H. S.; Zhan, X.; Liu, Y.; Zhu, D.; Tang, B. Z., *Chem. Commun.* **2001**, 1740-1741.

14. Jelley, E. E., *Nature* **1936,** *138*, 1009-1010.

15. Spano, F. C.; Silva, C., *Annu. Rev. Phys. Chem.* **2014,** *65*, 477-500.

16. Lill, R., *Nature* **2009,** *460*, 831-838.

17. Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Lidwig, M. L., *Science* **1994,** *266*, 1669-1674.

18. Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J., *Nature* **2009,** *460*, 823-830.

19. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M., *J. Biol. Inorg. Chem.* **2008,** *13*, 1205-1218.

20. Wächtershäuser, G., Chemoautotrophic Origin of Life: The Iron–Sulfur World Hypothesis. In *Geomicrobiology: Molecular and Environmental Perspective*, Barton, L. L.; Mandl, M.; Loy, A., Eds. Springer Netherlands: Dordrecht, 2010; pp 1-35.

21. Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K., *Annu. Rev. Biochem.* **2005,** *74*, 247-281.

22. Dos Santos, P. C.; Johnson, D. C.; Ragle, B. E.; Unciuleac, M. C.; Dean, D. R., *J. Bacteriol.* **2007,** *189*, 2854-2862.

23. Malkin, R.; Rabinowitz, J. C., *Biochem. Biophys. Res. Commun.* **1966,** *23*, 822-827.

24. Pandelia, M. E.; Lanz, N. D.; Booker, S. J.; Krebs, C., *Biochim. Biophys. Acta* **2015,** *1853*, 1395-1405.

25. (a) Johnson, C. E.; Bray, R. C.; Cammack, R.; Hall, D. O., *Proc. Natl. Acad. Sci. USA* **1969,** *63*, 1234-1238; (b) Rouault, T. A., *Nat. Chem. Biol.* **2015,** *11*, 442-445.

26. Hanson, G.; Berliner, L.; SpringerLink, Metals in Biology Applications of High-Resolution EPR to Metalloenzymes. In *Biological Magnetic Resonance*, [Online] Springer New York,: New York, NY, 2010; pp. XIX, 419p.

27. Que, L., *Physical Methods in Bioinorganic Chemistry: Spectroscopy and Magnetism*. University Science Books: Sausalito, Calif., 2000.

28. Schrauzer, G. N.; Deutsch, E., *J. Am. Chem. Soc.* **1969,** *91*, 3341-3350.

29. (a) Kadish, K. M.; Smith, K. M.; Guilard, R., *The Porphyrin Handbook*. Academic Press: San Diego, 2000; (b) Bertrand, E. M.; Allen, A. E.; Dupont, C. L.; Norden-Krichmar, T. M.; Bai, J.; Valas, R. E.; Saito, M. A., *Proc. Natl. Acad. Sci. USA* **2012,** *109*, E1762-1771.

30. Banerjee, R.; Ragsdale, S. W., *Annu. Rev. Biochem.* **2003,** *72*, 209-247.

31. Marsh, E. N., *Essays Biochem.* **1999,** *34*, 139-154.

32. Jarrett, J. T.; Huang, S.; Matthews, R. G., *Biochem.* **1998,** *37*, 5372-5382.

33. Evans, J. C.; Huddler, D. P.; Hilgers, M. T.; Romanchuk, G.; Matthews, R. G.; Ludwig, M. L., *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 3729-3736.

34. Gonzalez, J. C.; Banerjee, R. V.; Huang, S.; Sumner, J. S.; Matthews, R. G., *Biochem.* **1992,** *31*, 6045-6056.

35. Pejchal, R.; Ludwig, M. L., *PLoS Biol.* **2005,** *3*, e31.

36. Finke, R. G.; Hay, B. P., *Inorg. Chem.* **1984,** *23*, 3041-3043.

37. Ochiai, E. I., *Metal Ions in Biological Systems, Vol 30* **1994,** *30*, 255-278.

CHAPTER 2: SUPPRESSION OF KASHA'S RULE: A MECHANISM TO DESCRIBE THE ANOMALOUS EMISSION OF BORON DIFLUOROHYDRAZONE DYES

Material from this dissertation chapter has been published in the following form:

Qian, H., Cousins, M. E., Horak, E. H., Wakefield, A., Liptak, M. D., Aprahamian, I.. (2017). Suppression of Kasha's rule as a mechanism for fluorescent molecular rotors and aggregation-induced emission. *Nat. Chem.*, **9**, 83-97.

SECTION 2.1. INTRODUCTION

General structure of boron difluorohydrazone dyes. A family of organic dyes, derivatives of boron difluorohydrazone (BODIHY), have been synthesized by our collaborators (Professor Ivan Aprahamian and Hai Qian, Dartmouth College, Department of Chemistry) (**Fig. 2.1**).

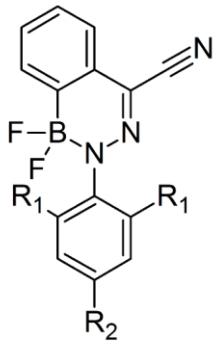


Figure 2.1. The general structure of the BODIHY derivatives primarily discussed in this chapter.

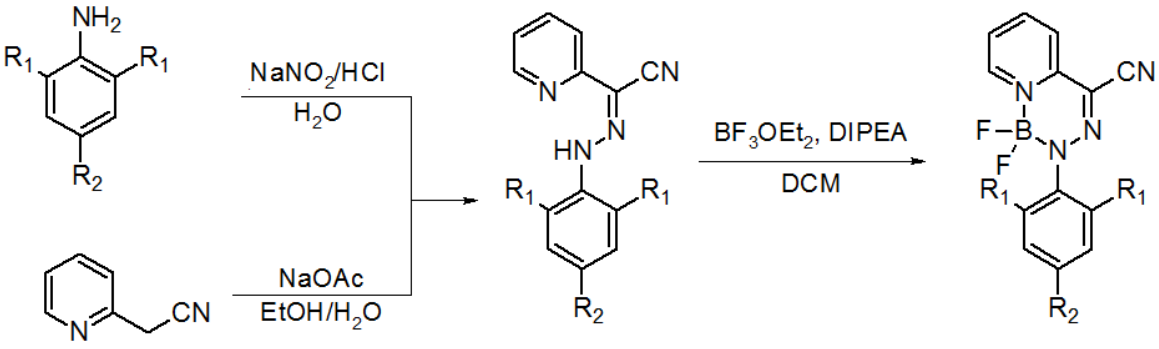
The BODIHY dyes are composed of two conjugated fluorophores: the boron containing conjugated heterocycle and the pendant phenyl ring (**Fig. 2.1**). Derivatives have been synthesized where the *ortho*- (compound 2) or *para*- positions of the pendant phenyl are substituted with electron withdrawing (EW) groups, such as in compounds 3-5, or electron donating (ED) groups, as seen in compounds 6-8 (**Table 2.1**).

Table 2.1. BODIHY dyes discussed in this chapter. R_1 and R_2 refer to the annotation of **Fig. 2.1**.

BODIHY	*ortho* substitutions (R_1)	*para* substitution (R_2)
1	-H	-H
2	$-CH_3$	-H
3	-H	-CN
4	-H	$-NO_2$
5	-H	$-COOCH_3$
6	-H	$-OCH_3$
7	-H	$-NH(CO)CH_3$
8	-H	$-N(CH_3)_2$

Three representative compounds will be discussed in more detail throughout this chapter: the parent BODIHY, compound 1; a representative EW BODIHY, compound 4 (*para*-nitro); and a representative ED BODIHY, compound 6 (*para*-methoxy). The addition of ED and EW groups offers a direct synthetic route to tune the brightness and viscosity dependence of the BODIHY derivative and will be discussed later in this dissertation.

Generalized synthesis. The synthesis of this fluorescent molecular rotor (FMR) was first published in 2012 (**Scheme 2.1**).[1] The diazonium salt was formed from a substituted aniline substrate. Tunable synthesis of BODIHY derivatives is achieved by synthesizing unique aniline substrates in the first step of this synthetic pathway. To synthesize the hydrazone product (the BODIHY precursor), a condensation reaction of the diazonium salt and 2-pyridineacetonitrile with sodium acetate was performed. The hydrazone precursor was incubated overnight with boron trifluoride diethyl etherate to produce the BODIHY species.



Scheme 2.1. The synthesis of the BODIHY complexes discussed in this chapter.

Experimental observations indicated that the BODIHY dyes are fluorescent molecular rotors. The family of BODHIY dyes are a series of blue-light emitting fluorophores that are attractive for many applications because of their viscosity sensitivity, brightness, as well as large Stokes shifts. The intensity of the emitted fluorescence from BODIHY family is enhanced when in the solid state.[1] The phenomenon responsible for this enhancement is by aggregation-induced emission (AIE). Viscosity dependence is a hallmark trait of FMRs, an up-and-coming class of fluorescent molecules.[2] This

phenomenon has been elusive to many organic dyes; in fact, most dyes undergo aggregation caused quenching (ACQ) in the solid state regardless of any attractive photophysical properties when in solution.[3] The ACQ phenomenon is far more common than AIE and introduces many challenges when the luminogens are used for technological applications.[4]

The most desirable attributes for fluorophores are viscosity sensitivity, brightness, as well as large Stokes shifts. Dyes with these impressive parameters have potential in material applications such as organic light emitting diodes (OLEDs) and microscopic viscosity sensors. Some challenges hindering the advancement of methodologies to produce more efficient luminogens include interference from polarity sensitivity (spectral interference), small Stokes shifts (self-quenching), and decreased brightness in the solid state (low quantum yield). The most promising BODIHY derivative, compound 4, is superior to the commercially available molecular rotor, 9-(dicyanovinyl) julolidine (DCVJ), by measured viscosity sensitivity, brightness, and Stokes shift.[5,6]

Aggregation influences the emission of fluorophores. The aggregation of fluorophores is the interaction of one or more molecules as either fluorophore-fluorophore or fluorophore-solvent interactions. For aggregation caused quenching, the interaction between two fluorophores often prevents radiative relaxation pathways or allows non-radiative pathways to become more competitive. In aggregation induced emission, the radiative pathways to become more competitive than the non-radiative relaxation pathways.

ACQ was first reported in the 1950's and the mechanism has been largely agreed upon: the conformation of the fluorophores in the aggregate state leads to the rate of non-radiative relaxation from the excited state to be more competitive than radiative relaxation (fluorescence or phosphorescence).[7,8] Some conformations that lead to ACQ include the formation of weak eximers or π-π interactions.[7,8] A common approach used to prevent ACQ is to decorate the fluorophore with bulky groups that interrupt fluorophore-fluorophore interactions. When developing OLEDs, ACQ is especially deleterious as the luminescent molecules are often required to be in the solid state for industrial applications. This results in a drastic decrease in the quantum yield of the material.

AIE was first reported in the literature in 2001; very few mechanisms for AIE have been proposed, ergo the quantum mechanical origin of AIE is not well understood as compared to ACQ.[9] One proposal is that AIE arises from the restriction of intramolecular motions (RIM).[10] The RIM phenomenon is the combination of other AIE phenomena: restriction of intramolecular rotations (RIR) and restriction of intramolecular vibrations (RIV).[10] The RIM phenomenon proposed that any motion of the fluorophore will to curtail radiative relaxation pathways by consuming the excitation energy. The RIM phenomenon assumes that the energy of molecular motion is proportional to the *amplitude* of the rotation or vibration. This proposal violates the Born-Oppenheimer approximation (a building block of quantum mechanical theory). Instead the Born-Oppenheimer approximation asserts that the energy is proportional to *frequency* of the molecular motion.[11] Instead of a quantum mechanical mechanism that can be used to describe the photophysics of dyes, the RIM phenomenon is simply the accumulation of experimental observations. These

phenomena cannot be used to describe the photophysical mechanism behind the enhanced emission of AIE fluorophores.

Instead of restriction of the motion of the fluorophore leading to of a consumption of energy as the RIM phenomena describe, aggregation must instead prevent a conformation that avoids an increase in favorability of any non-radiative relaxation pathway. One such mechanism that can explain the emission enhancement of dyes in the aggregate state is the twisted intramolecular charge transfer (TICT) mechanism.

TICT fluorophores are composed of three conjugated subunits: electron donating moiety, an electron-rich linker, and the electron accepting moiety. After photoexcitation, a charge transfer (CT) occurs from the electron-donating unit to the electron-accepting unit while the molecule is in the planar locally excited (LE) conformation. The complex then undergoes a twisting motion to form a lower energy TICT complex. Some TICT fluorophores exhibit dual S_1 emission arising from the two conformers: LE and twisted conformation.[12] Other fluorophores, such as DCVJ, only exhibit emission from the LE S_1 state. The TICT S_1-S_0 energy gap is significantly smaller than the LE energy gap and the favored pathway is a non-radiative one.[13] DCVJ is an example of a single emission TICT fluorophore (**Fig 2.2**).

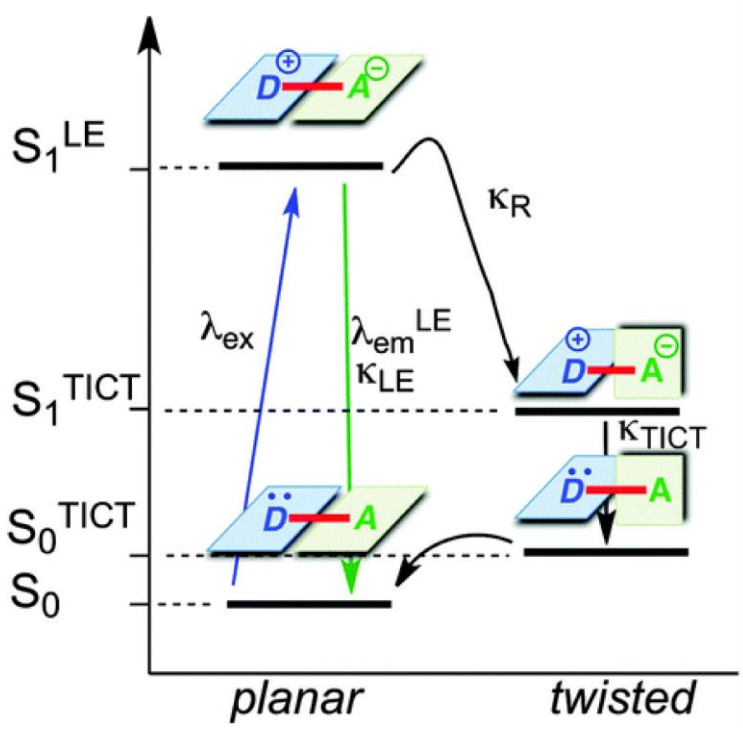


Figure 2.2. Jablonski diagram of a single emission TICT mechanistic pathway for DCVJ. Blue arrow indicates photoexcitation, green arrow indicates radiative relaxation from S_1, and the black arrows indicate non-radiative relaxation processes. The blue parallelogram ("D") represents the electron-donor, green parallelogram ("A") represents the electron-donor, and the connecting red line represents the electron rich linker. Following the photoexcitation with λ_{ex}, fluorescent relaxation occurs at λ_{em}^{LE} with a deexcitation rate of κ_{LE} from the locally-excited planar state (LE). The formation intramolecular charge transfer state is followed by a rotation around the dicyanovinyl group. Radiative relaxation only occurs from S_1^{LE} because the S_1-S_0 energy gap in the LE state is three-times as large as the energy gap in the twisted conformation. Figure taken from reference 12.

In the initial study of the BODIHY dyes, it was shown that the emission of compound 1 was enhanced in the solid state through AIE.[1] At the time, it was proposed that this enhanced emission was due to RIR. In accordance with the RIR phenomenon, the rotation of the phenyl ring would be hindered in the solid state as compared to rotation in the solution phase.[1] By restricting the free rotation of this pendant phenyl, it was presumed that many non-radiative relaxation pathways would be prevented which led to emission enhancement. Again, RIR is not a mechanism but simply the combination of many experimental observations, consequently the RIR phenomenon cannot be used to describe the emission of the BODIHY dyes. For many FMRs, the TICT mechanism can be used to explain the observed AIE and is easily validated experimentally. If a TICT mechanism is responsible for the observed AIE, the solvent will have a significant influence on the formation of this complex. Relaxation from the lower energy twisted state is promoted in polar solvents as hydrogen bond formation between the solvent and the TICT state is a stabilizing force. The stabilization of the TICT complex of a single emission fluorophore would lead to a significant decrease in intensity corresponding to favored relaxation from the TICT complex (non-radiative).[6] The stabilization of the TICT complex of a double emission fluorophore would lead to the red-shifting of the emission maximum corresponding to radiative relaxation from the lower energy TICT complex.[12]

A better understanding of the quantum mechanical origins of the AIE mechanism would improve the rational design of new luminogens for a myriad of applications and fields. Increased insight into the photophysical mechanism responsible for the viscosity

sensitive emission of the BODIHY dyes was achieved through computational and spectroscopic means.

SECTION 2.2. METHODS

Viscosity sensitivity of BODIHY emission. To better understand the influence on solvent viscosity on emission enhancement, a series of solvent cocktails of increasing viscosity were used to restrict rotation (**Table 2.2**).

Table 2.2. Solvent cocktails of increasing viscosity.

Ethylene glycol (%)	Glycerol (%)	Viscosity (cP)
20	80	620.7
40	60	283.1
60	40	124.0
70	30	80.8
80	20	52.0
100	0	20.8

This viscosity gradient is commonly used to investigate the viscosity sensitivity of molecular rotors as changes in viscosity are significant while only small changes in polarity occur.[2] Due to limitations of solubility, the BODIHY dyes were first dissolved in DMSO. This was used to stain ethylene glycol (EG); each solvent cocktail contains 10% of this stained EG solution. The viscosity of each mixture was determined with the following equation to consider the weight fraction of the solvent mixture (**Eqn. 2.1**).

$$\ln \eta_{mix} = \sum_{i=1}^{n} w_i \cdot \ln \eta_i \qquad \textbf{(Equation 2.1)}$$

Here, η_{mix} and η_i represent the viscosity of the solvent cocktail and the components and w_i represent the weight fraction of the component. To find the weight fraction of each

solvent, the density values used for DMSO, EG and glycerol (Gly) were 1.10, 1.11, and 1.26 g/mL respectively.

Polarity dependence of BODIHY emission. As the TICT mechanism describes the formation of a CT complex, the influence of a solvent polarity on the emission of the BODIHY dyes was also investigated. The polarity dependence of the BODIHY dyes' emission was explored using a binary solvent cocktail of water and dioxane (**Table 2.3**).

Table 2.3. Solvent cocktails of increasing polarity.

Dioxane (%)	Water (%)	$\varepsilon_{mixture}$
100	0	2.30
90	10	9.88
80	20	17.50
70	30	25.16
60	40	32.86
50	50	40.62
40	60	48.42
30	70	56.27
20	80	64.17

The dielectric constants were used to quantify the polarity of these solvent mixtures: ε_{water} = 80.1 and $\varepsilon_{dioxane}$ = 2.3. The viscosities of the component of the cocktail are both approximately 1 cP. Any influence of solvent viscosity on emission intensity should be minimal and this experiment will provide information solely on the polarity sensitivity of the BODIHY dyes.

DFT Calculations. The photophysics of the BODIHY dyes were explored through modeling by density functional theory (DFT) and time-dependent DFT (TDDFT) calculations. The computations were performed on the IBM Bluemoon cluster at the Vermont Advanced Computing Core. All computations to predict the electronic states of the BODIHY dyes (compounds 1-4, 6 and 8) used model structures prepared in ArgusLab (Planaria Software); structures were modified from the crystal structure of compound 1 (CCDC number 841443). This structure was then optimized using DFT geometry optimization on the ORCA 3.0.0 software package.[14]

Three functionals were tested to identify the best fit for the BODIHY dyes. Ultimately, the electronic structures of the dyes were modeled using the PBE functional paired with a triple-zeta basis set with a tight SCF convergence criteria.[15,16] Experimental verification of predicted parameters is the best manner in which to identify if a chosen functional is a good fit to the system it is attempting to model. The functional that most accurately predicted the molar absorptivity (ε) was selected as the best fit for modeling the BODIHY dyes. The predicted intensity of the Abs transitions is rooted in the shape of the donor and acceptor molecular orbitals. The predicted energy of the Abs transition arises from the energy difference between the occupied and unoccupied molecular orbitals. Typically, predicted energies are overestimated or underestimated by a constant; this would not impact the calculations pertaining to the molecular orbitals. The more significant parameter is the intensity as the modeling required to better understand the AIE of the BODIHY dyes pertains to the valence orbitals not the energy difference between the highest occupied molecular orbital and lowest unoccupied molecular orbital.

DFT models of compounds 1-4, 6, and 8 as a function of rotor angle were developed by a series of constrained geometry optimizations. For each constrained geometry optimization, the dihedral angle of the molecular rotor was held constant, while all other bond lengths and angles could fully relax. 36 different geometries were generated for each of the six species where the rotor dihedral angle was varied in 10° increments. (**Fig. 2.3**).

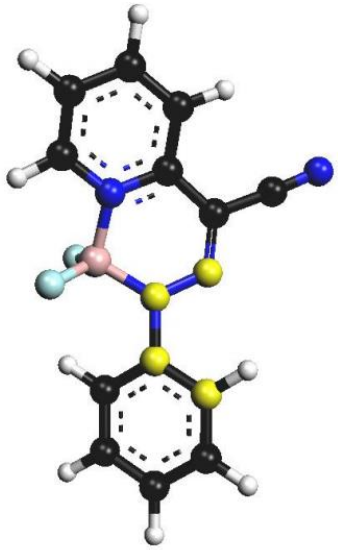


Figure 2.3. General structure of BODIHY used for geometry optimizations for compounds 1-4, 6 and 8 generated in ArgusLab (Planaria Software). Black spheres represent carbon, blue spheres represent nitrogen, pink spheres represent boron, light blue spheres represent fluorine and white spheres represent hydrogen atoms. The highlighted N–N–C–C (yellow atoms) dihedral angle was varied in 10° increments.

The total SCF energy of each constrained geometry optimization was used to assess the ground state electronic energy of each species as a function of rotor angle. The structures generated by these DFT calculations were employed for TDDFT modelling of

the excited state electronic structure. The TDDFT calculations for compounds 1-4, 6, and 8 as a function of rotor angle used the same density functional and basis set as described above. For each structure, TDDFT was used to calculate the excitation energies and intensities of transitions from the electronic ground state to the first 20 electronic excited states within an expansion space of 120 vectors. In addition, TDDFT was employed for an S1 excited state geometry optimization of compound 1 starting from its minimum on the rotor potential energy surface (PES). To aid analysis of the TDDFT data, TDDFT difference plots were generated by ORCA, and plotted in gOpenMol with isodensity values of ±0.003 a.u.. The Abs spectra of compounds 1-4, 6, and 8 were simulated based upon the TDDFT data by convoluting Gaussian-shaped bands with full width at half maximum bandwidths of 4000 cm^{-1}.

Circular dichroism (CD) spectroscopy. Compounds 1, 4 and 6 were prepared for cryogenic temperature CD spectroscopy in the most viscous solvent cocktail (80% Gly/ 20% EG (*v/v*)) before being loaded into a custom-built copper/quartz sample cell. The copper/quartz sample cell was flash frozen in liquid nitrogen prior to loading the sample into the Oxford Spectromag. 77 K CD data were acquired on a home-built setup consisting of a Jasco J-815 spectropolarimeter, an Oxford SM4000-8T Spectromag, and a Mercury iTC temperature controller. CD data were acquired in the 600-300 nm range with a scanning speed of 200 nm/min, a bandwidth of 1 nm, a digital integration time of 0.25 s, and a data pitch of 0.5 nm. The spectral baseline arising from light scattering off the frozen sample was corrected by a curve with λ^{-4} dependence.

Excitation-dependent emission spectroscopy. Compounds 1, 4 and 6 were dissolved in dichloromethane then placed on a MaxiMix Vortex Mixer for a minimum of one min to ensure sample homogeneity. The fluorescence spectra were collected immediately after sample was prepared. Fluorescence spectra were collected on a Photon Technology International QuantaMaster 4 spectrofluorometer outfitted with a LPS-100 lamp power supply, ASOC-10 electronics interface, MD-4 motor driver control, and a Model 814 photomultiplier detector system. Data were acquired with a step size of 0.5 nm, an integration time of 2 s, and slit widths of 7.8 nm. Samples were excited at 420 nm and the corresponding emission spectrum was collected from 430-650 nm. When the hydrazones were excited by a lower energy wavelength, the emission spectra were collected from (λ_{ex} + 15 nm) to 650 nm. An emission spectrum was collected of dichloromethane using identical parameters to identify features arising from Raman scattering of the solvent. The presented data were processed by subtracting the corresponding spectra of dichloromethane to remove features aring from Raman scattering of the solvent.

SECTION 2.3. RESULTS

Excitation wavelengths selected to excite the dyes were determined from the peak maximum in the UV-vis absorption (Abs) spectrum of the BODIHY derivative in EG (**Fig 2.4**). For compound 1, the emission intensity increased by almost 500 a.u. (over 600% increase) when comparing the least viscous cocktail (100% EG) to the most viscous solvent cocktail (80% EG, 20% Gly) (**Fig. 2.4**).

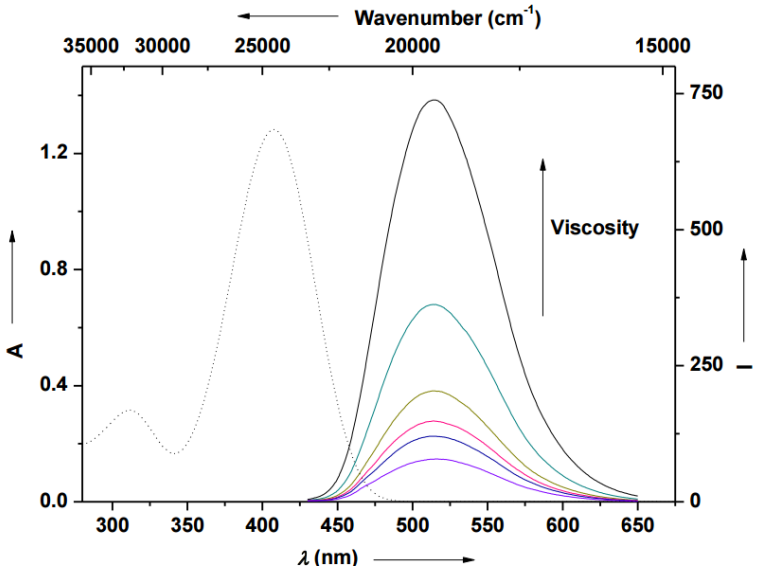


Figure 2.4. Compound 1 (10 μM) in EG-Gly cocktails of increasing viscosities. The dashed line represents an Abs spectrum and the solid lines indicate the emission spectra. The excitation slit widths were 2.5 nm and the emission slit widths were 5.0 nm.

Based on the positive value for χ for the compounds 1-8, all BODIHY dyes exhibit AIE (**Table 2.4**). Both viscosity sensitivity and overall brightness were determined from regression lines from the double logarithmic plots of emission intensity compared to viscosity for each BODIHY dye.[5] The slope gives the value for viscosity sensitivity and brightness is extracted from this double-logarithmic plot as $a = 10^b$. Overall, the EW BODIHY dyes (compounds 3-5) exhibit higher sensitivity to the viscosity of the solvent compared to the ED BODIHY dyes. With an average Stokes shift greater than 5,000 cm^{-1}, the family of BODIHY dyes are unlikely to self-quench.[17] The *para*-substituent offers a direct synthetic route to decrease the overlapping of the emission and excitation bands.[5]

Table 2.4. The photophysical properties of the BODIHY dyes (See **Table 2.1** to clarify the substituents on the phenyl).

BODIHY	Abs λ_{max} (nm)/ ε (M^{-1} cm^{-1})	λ_{em} (nm)/ λ_{ex} (nm)	Stokes Shift (nm)/(cm^{-1})	Viscosity sensitivity (χ)	Overall brightness (a)
1	408/ 25640	514/ 412	106/ 5050	0.66	9
2	371/ 16360	488/ 370	117/ 6460	0.49	4
3	406/ 30180	500/ 405	94/ 4630	0.68	29
4	415/ 68120	493/ 417	78/ 3810	0.64	110
5	409/ 19380	505/ 408	96/ 4650	0.69	19
6	426/ 32580	551/ 421	125/ 5330	0.54	24
7	425/ 19420	552/ 420	127/ 5410	0.49	34
8	484/ 58920	636/ 485	152/ 4940	0.32	2

The polarity sensitivity for compounds 1, 4, and 6 were explored using these solvent mixtures of increasing polarities. There was minimal solvatochroism shown in the Abs spectra of the BODIHY dyes.[5] There was also little influence of polarity on the emission intensity of the BODIHY dyes (**Fig. 2.5**).

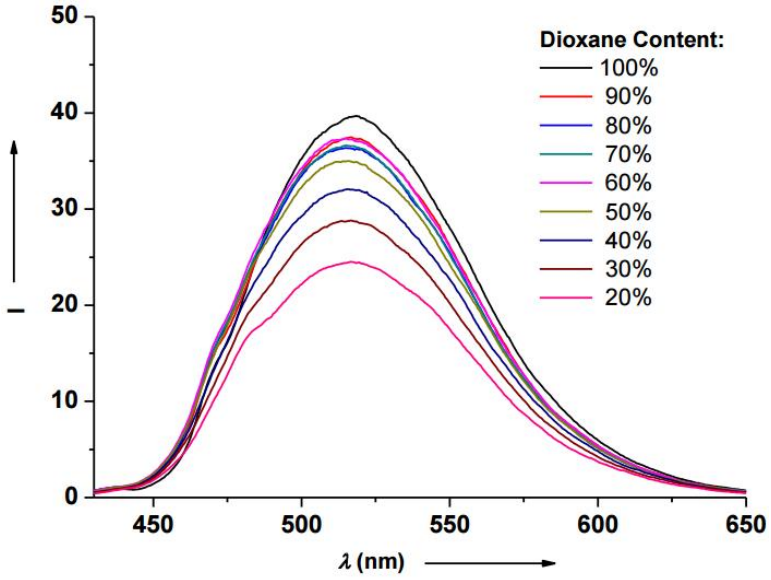


Figure 2.5. Compound 1 (10 μM) in binary solvent cocktails (water and dioxane) of increasing polarity. The excitation slit widths were 2.5 nm and the emission slit widths were 5.0 nm.

For compound 1, the emission intensity decreased by approximately 15 a.u. (40% decrease) when comparing the emission spectrum of BODIHY in the most polar cocktail to the least polar solvent cocktail (**Fig. 2.5**). Similar results were observed for compounds 4 and 6.[5] Thus, changes in solvent polarity are not responsible for the emission enhancement in the solid state; there was little change in the energy of the emission maxima for compounds 1, 4, and 6. Therefore, a CT mechanism was not responsible for the AIE of

the dyes. Of the BODIHY dyes, compound 4 had the largest shift in emission maxima of the most and least polar solvent cocktail (450 cm^{-1}).[5] For reference, the shift in the emission maximum for a TICT fluorophore, DCVJ, is over 1000 cm^{-1} under the same conditions.[6] A significant bathochromic shift of the emission maxima in polar solvents would suggest a CT-based mechanism.

Thus, the solvent polarity only slightly influences the emission maxima and intensity, while the microenvironment viscosity is highly-influential on the emission intensity of the BODIHY dyes. As such, the mechanism responsible for the AIE of the BODIHY family cannot be a CT mechanism such as TICT. In order to better understand the viscosity sensitivity of the BODIHY dyes' emission, DFT and TDDFT were used to investigate the electronic transitions related to the BODIHY emission. A photophysical mechanism responsible for the AIE of the BODHY dyes was determined through the DFT and TDDFT modeling and corroborated with experimental data.

PBE was selected to model the BODIHY dyes because it most accurately predicted the molar absorptivity. The PBE functional was carefully chosen for the DFT and TDDFT modelling discussed in this chapter. PBE is a simple, generalized gradient approximation (GGA) functional that is applicable to a variety of systems.[15] The PBE predicted parameters of the absorption spectra of compounds 1-4, 6, and 8 are presented in **Table 2.5**.

Table 2.5 Photophysical properties of select BODIHY dyes as determined experimentally and computationally.

BODIHY	Abs λ_{max}[a] (nm)	Abs λ_{max}[b] (nm)	ε[a] (M^{-1} cm^{-1})	ε[b] (M^{-1} cm^{-1})	G.S. barrier (kcal/mol)	E.S. barrier (kcal/mol)
1	408	403	25640	28040	4.1	7.1
2	371	329	16360	20320	16.5	16.0
3	406	414	20180	28600	5.0	9.8
4	415	414	68120	40840	5.1	7.7
6	426	469	32580	26170	4.4	5.9
8	484	530	19420	22560	5.3	3.9

[a] Experimental values (in EG)
[b] PBE TDDFT-predicted values

Other functionals were used to model the photophysics of these BODIHY dyes and will be discussed later in this chapter. The PBE functional was selected because of the good agreement between the wavelengths and intensities of the TDDFT-predicted and experimental Abs features (**Table 2.5**). For example, PBE correctly predicted the red-shift from compound 1 to 4 (coinciding with the addition of the *para*-nitro group on the phenyl pendant), as well as the blue-shift from compound 1 to 2 (the addition of two *ortho*-methyl groups on the phenyl pendant).[5] The TDDFT-predicted Abs spectra were simulated by

convoluting Gaussian-shaped bands with full width at half maximum bandwidths of 4000 cm^{-1}; this bandwidth was derived from the spectrum of compound 1 in EG (**Fig. 2.6**).

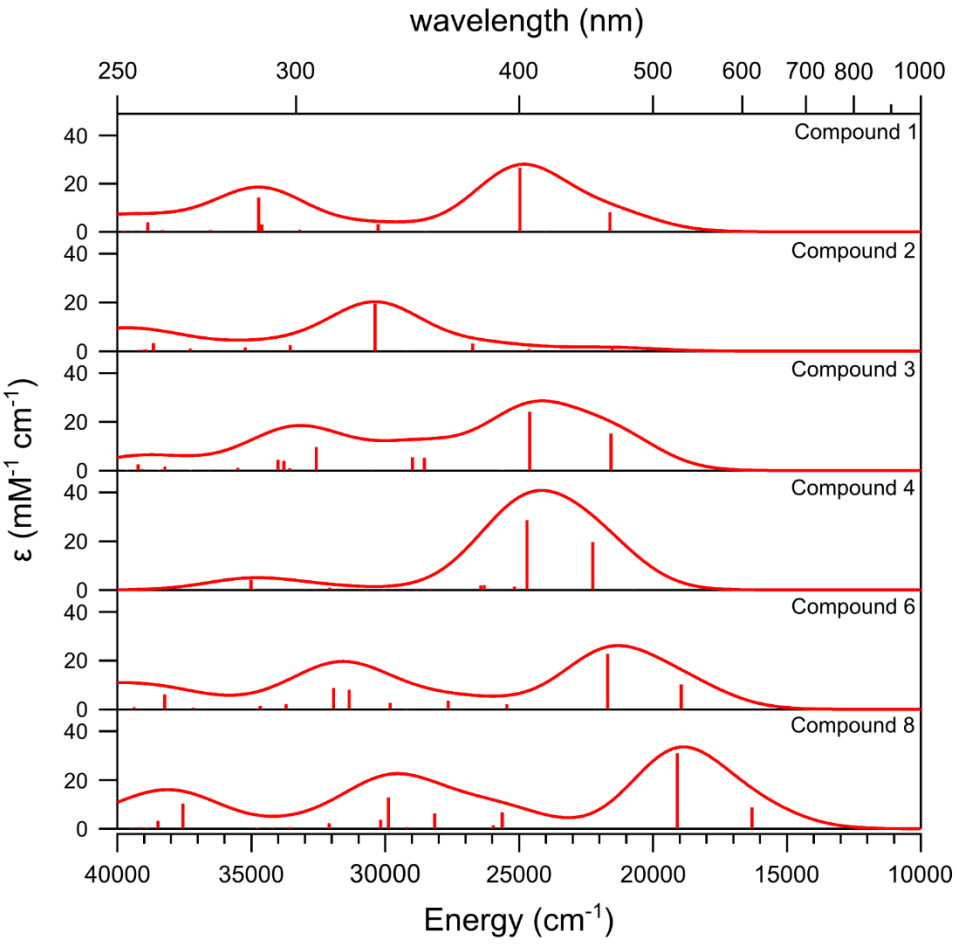


Figure 2.6. PBE TDDFT-predicted absorption spectra of compounds 1-4, 6, and 8. Here the vertical sticks represent the TDDFT-predicted transition energies and intensities of unique features while the spectral traces arose from the convolution of Gaussian-shaped bands with full width at half maximum bandwidths of 4000 cm^{-1}.

Because of the good agreement between the experimental and predicted values, it can be assumed with confidence that properties extracted from these calculations arise from accurate structural and electronic origins. Thus, the photophysical properties of the

BODIHY dyes could be further explored by both DFT and TDDFT. To understand the origins of the viscosity dependent emission, the ground and excited electronic states were investigated by DFT and TDDFT. **Figure 2.7** illustrates the dependence of the brightness of emission on the dihedral angle of the rotor; brightness is indicated by the oscillator strength (f_{osc}) and darker red shading of the diamond is indicative of a brighter emission.

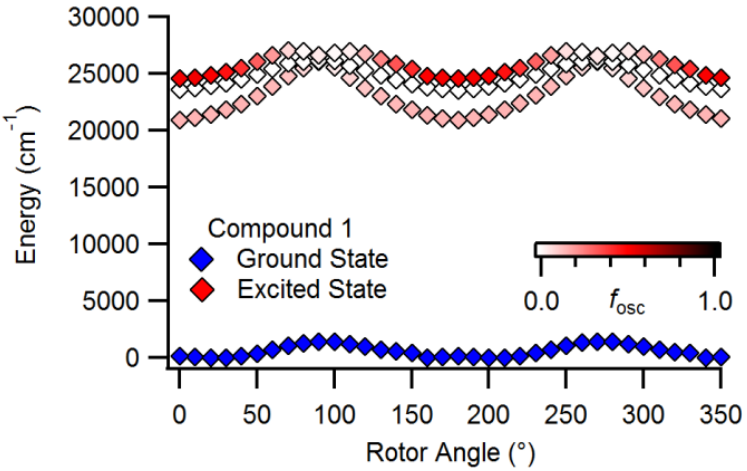


Figure 2.7. PBE TDDFT-predicted ground (blue diamonds) and excited states (shaded red diamonds) of compound 1. The red shading of the excited state is based upon the TDDFT-predicted oscillator strength for an electric dipole-allowed transition between the ground and excited state. Only electric dipole-allowed transitions were screened.

For compound 1 (as well as compounds 3, 4, 6, and 8), the emissive state is not the lowest energy excited state, S_1, but instead a higher lying excited state! This observation is contrary to Kasha's Rule which states that excitation into higher lying excited states results in internal conversion (IC) to the lowest lying excited state, S_1, followed by radiative relaxation to the ground state, S_0.[18] Instead, S_1 is a *dark state* as implied by the weak

oscillator strength; relaxation from S_1 will be predominantly non-radiative. From the PES, it can be concluded that enhanced emission is directly related to the conformation of the molecule (**Fig. 2.7**).

Other functionals fail to reflect experimental data. The PBE predicted spectra of BODIHY dyes are in good agreement with the experimentally derived absorption spectra. This indicated that the electronic and structural origins from which all DFT- and TDDFT-predicted parameters would be derived was accurate. The other functionals screened, PBE0 and ωB97X, are more computationally costly than PBE and differ in their approach for approximating the E_{XC}. In terms of computational costs, calculating the TDDFT-predicted excited state as a function of the dihedral angle (36 calculations per compound to model a full rotation of the phenyl around the heterocycle), rapidly became exceedingly costly in terms of computer resources and time. However, if either functional could model the photophysical properties with superior accuracy as compared to PBE, that increase in required computer time would be well invested.

PBE0 was not used to model BODIHY because it overestimated the intensity of the lowest energy feature. The more moderately advanced functional, PBE0, was screened but ultimately deemed inferior to PBE. PBE0 is related to the PBE functional; the PBE method for determining the correlation energy is retained but the PBE0 exchange energy is the summation of the Hartree-Fock exchange energy and the PBE exchange energy.[19] PBE0 predicted a similar absorption spectra as PBE; the PBE0-predicted energy of the main transition of compound 1 was more accurate (**Fig. 2.8** and **Table 2.3**).

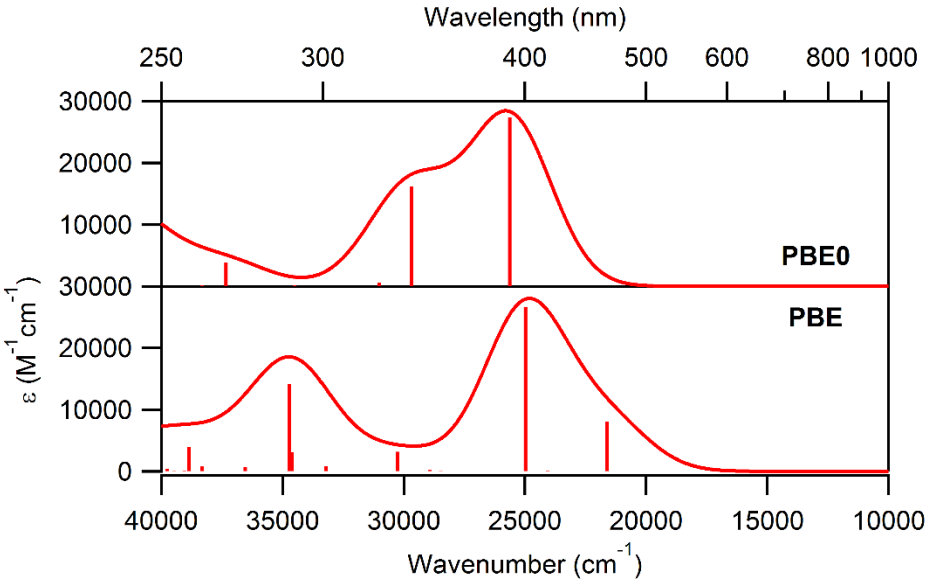


Figure 2.8. The PBE0 (top) and PBE (bottom) TDDFT-predicted absorption spectra for compound 1. The spectral curves arise from the convolution of Gaussian-shaped bands with full width at half maximum bandwidths of 4000 cm^{-1} and the vertical sticks are indicative of transition energies and intensities of unique features.

The PBE functional more accurately predicted the intensity of the main transition compared to PBE0 while the PBE0-predicted energy was more accurate. PBE0 predicted

that the BODIHY dye would not freely rotate in solution upon excitation which was not supported experimentally (**Table 2.6**). While inferior to PBE, the PBE0-predicted Abs spectrum was promised enough to model compound 1 by TDDFT using the PBE0 functional. The emissive state was predicted to be a third excited state, S_1, while the two excited states lower in energy, T_1 and T_2, were predicted to exhibit phosphorescence (**Fig. 2.9**).

Table 2.6. DFT and TDDFT predicted properties of compound 1: PBE and PBE0.

Functional	Abs λ_{max} (nm)	ε (M^{-1} cm^{-1})	G.S. barrier (kcal/mol)	E.S. barrier (kcal/mol)
PBE	403	28,000	4.1	7.1
PBE0	388	28,400	3.8	13.1

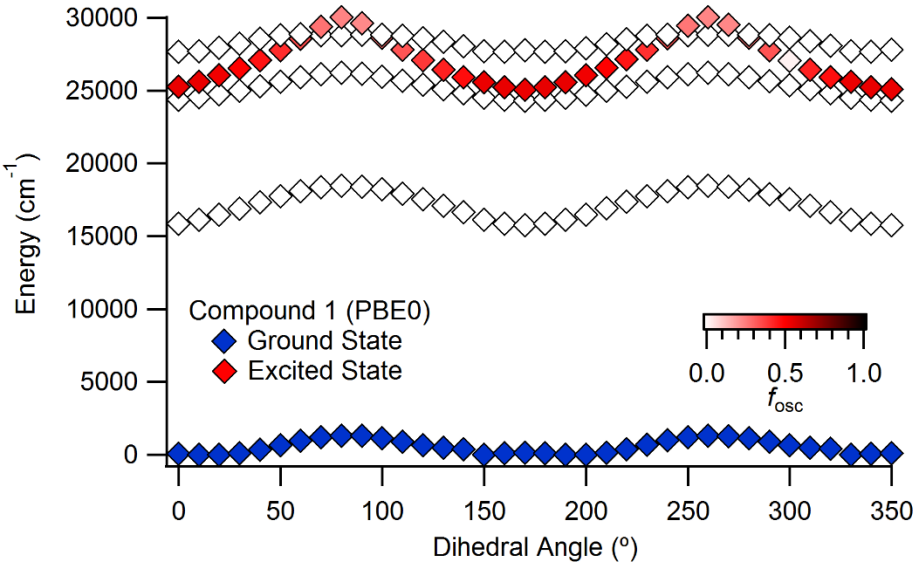


Figure 2.9. The PBE0 TDDFT-predicted PES. Here the ground state (blue diamonds) and excited states (shaded red diamonds) are presented as a function of rotor angle. The red shading of the excited state is based upon the TDDFT-predicted oscillator strength for an electric dipole-allowed transition between the ground and excited state. The unshaded excited state diamonds indicate possible electric dipole-not allowed transitions (phosphorescence). Therefore, the first two excited states would be T_1 and T_2 and the third excited state would be S_1.

The fluorescence lifetime data of compound 1 discourages the likelihood of phosphorescence.[5] Even though PBE0 more accurately predicted the energy of the main absorption feature, this functional led to a TDDFT-predicted model of the photophysical properties that was not corroborated by experimental data and was not chosen over the PBE functional.

ωB97X was not used to model the BODIHY dyes because it overestimates the intensity of the main feature by over 50% in the predicted Abs spectrum. ωB97X, a long-range hybrid functional, was also screened.[20] ωB97X does a poor job in replicating the experimental absorption spectrum of compound 1 (**Fig. 2.10**).

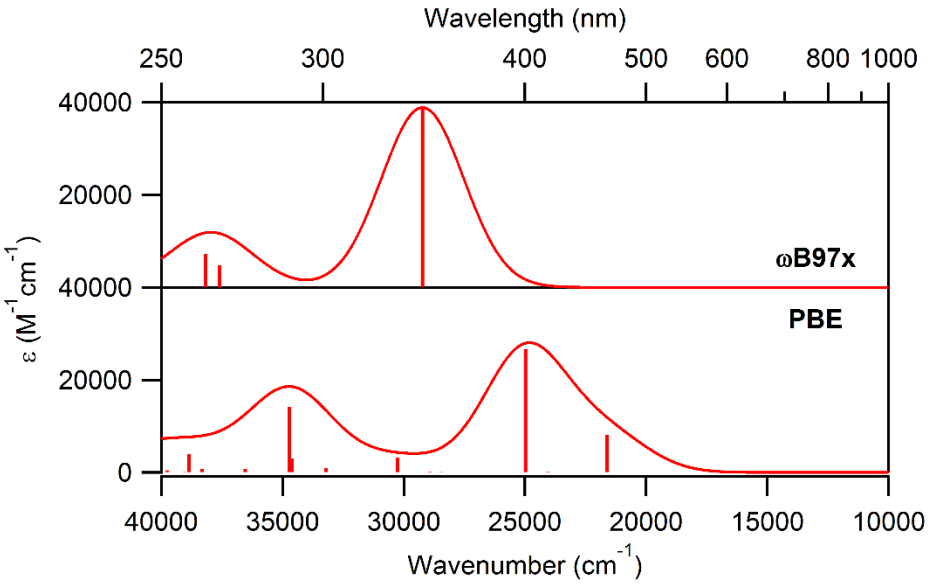


Figure 2.10. The ωB97X (top) and PBE (bottom) TDDFT-predicted absorption spectra for compound 1. Here the spectral curves arise from the convolution of Gaussian-shaped bands with full width at half maximum bandwidths of 4000 cm^{-1} and the vertical sticks are indicative of transition energies and intensities of unique features.

Another aspect that ωB97X severely miscalculates is the barrier to rotation in the excited state; a barrier of 14.3 kcal/mol implies the BODIHY would no longer be able to freely rotate in solution once excited. However, the most egregious error is that ωB97X severely overestimates the intensity of the main feature by over 50% **(Table 2.7**).

Table 2.7. DFT and TDDFT predicted properties of compound 1: PBE and ωB97X.

Functional	Abs λ_{max} (nm)	ε (M^{-1} cm^{-1})	G.S. barrier (kcal/mol)	E.S. barrier (kcal/mol)
PBE	403	28,000	4.1	7.1
ωB97X	342	38,800	2.7	14.3

ωB97X did not pass the initial test of predicting the absorption spectrum; consequently, it cannot be expected to accurately predict properties of the dyes. For this reason, ωB97X was not chosen to predict the myriad of photophysical parameters for the BODIHY family. The PES of compound 1 as predicted by ωB97X shows that the first excited state would be the emissive excited state (**Fig. 2.11**).

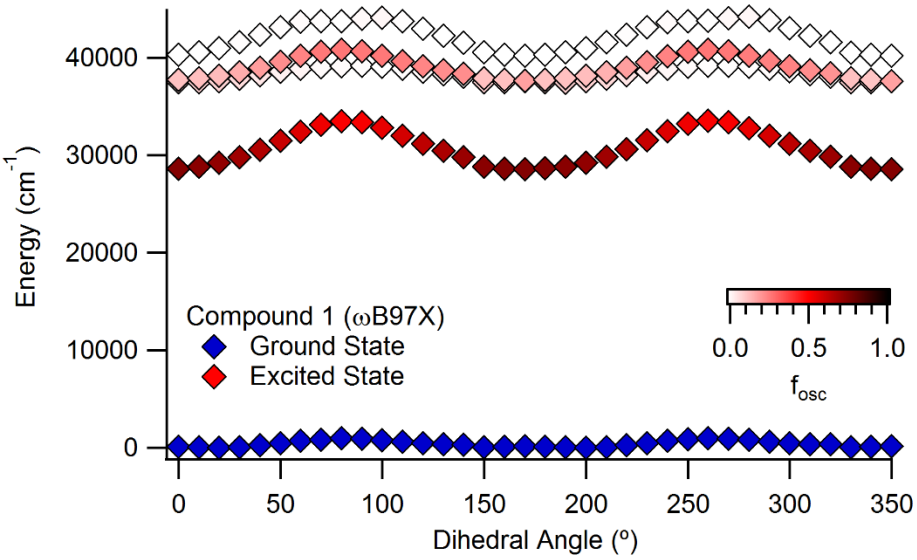


Figure 2.11. The ωB97X TDDFT-predicted PES. Here the ground state (blue diamonds) and excited states (shaded red diamonds) are presented as a function of rotor angle. The red shading of the excited state is based upon the TDDFT-predicted oscillator strength for an electric dipole-allowed transition between the ground and excited state.

The ωB97X-predicted PES predicts that the emission of compound 1 would follow Kasha's Rule: emission will occur from S_1.[18]Again, this functional produced a TDDFT-predicted model of the photophysical properties that was not corroborated by experimental data and was not chosen over the PBE functional.

Both hybrid functionals, PBE0 and ωB97X, poorly modelled the Abs spectra of the BODIHY dyes further cementing PBE as the proper choice as the functional for these FMRs. While anomalous emission, or emission from high-lying excited state, has been documented in the past, it is not common.[21] The SOKR mechanism describes a three state

system: the ground state (S_0), a weakly emissive or dark state (S_1), and the bright state ($S_{>1}$). A suite of spectroscopic experiments validated this novel three state mechanism.

Computational investigation of viscosity sensitivity. Here, a trend between viscosity sensitivity (**Table 2.3**) and barrier to rotation in the emissive state (**Table 2.5**) can be discerned. A lower barrier to rotation in the emissive state translates to a lower viscosity sensitivity. Compound 8 has the lowest viscosity sensitivity ($\chi = 0.32$) and lowest barrier to rotation (3.9 kcal/mol). Compounds 1 and 4 have similar barriers to rotation, 7.1 and 7.7 kcal/mol respectively, and viscosity sensitivities. Unsurprisingly, compound 2 does not fit this trend as it is not expected to behave as a FMR due to the high barrier to rotation (over 10 kcal/mol).

Computational evidence that a CT mechanism does not describe the enhanced emission of BODIHY in viscous environments. To better understand the nature of the excitation transitions for compounds 1-4, 6, and 8, the electronic structure of each excited state was modeled by TDDFT calculations. The optimized geometries from the DFT calculations were used as the equilibrium geometry of the dye at the ground state. The electronic transition from the donor orbital to the acceptor orbital is illustrated in the difference density plots. To better understand the subtle differences in the electronic transitions, difference density plots were produced from the TDDFT computations and optimized geometries of compounds 1-4, 6, and 8 (**Fig. 2.12**).

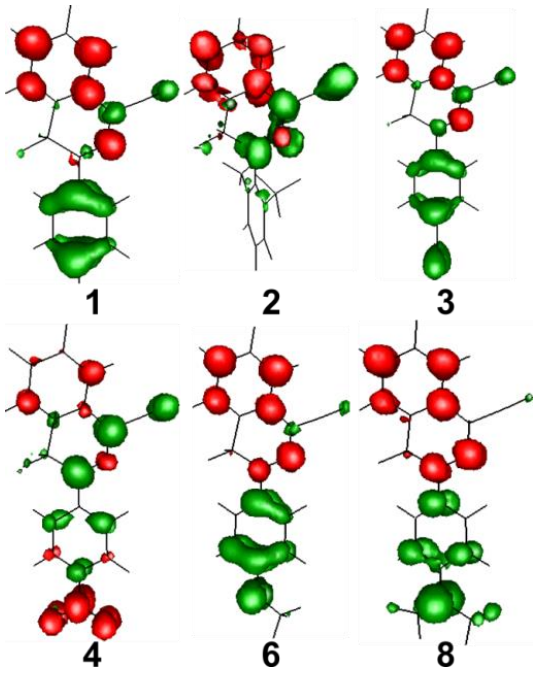


Figure 2.12. The difference density plots of compounds 1-4, 6, and 8 as predicted by TDDFT. The electronic structures were generated using gOpenMol and used isodensity values of ± 0.003 a.u.. The electron density is color coded in the following fashion: the green is indicative of loss of electron density while red indicates a gain of electron density.

Here the slight differences in the excitation transitions become more pronounced. These transitions all show the conjugated π-system of the fluorophore where the major transition can be described as $\pi \rightarrow \pi^*$ transition with some CT character. Some phenyl → heterocycle CT character contributes to the overall excitation transitions of compound 1, 3, 6, and 8. This could account for the minimal polarity sensitivity of the emission intensity of the BODIHY dyes. Compound 2 is unique in that the phenyl ring does not contribute to the excitation transition; instead, a $\pi \rightarrow \pi^*$ transition localized in the heterocycle dominates the molecule's excitation transition. Compound 4 also has a unique excitation transition.

The *para*-nitro group is involved in the excitation transition in a significant fashion: the BODIHY moiety is donating electron density into the nitro group. This explains why compound 4 has a high overall brightness which, at first glance, would be unexpected for this fluorophore. Usually a nitro group would quench the fluorescence through photoinduced electron transfer (PET).[22] The PET mechanism would lead to quenching because of a transition in electron density from the fluorophore into the nitro group during relaxation. This is circumvented in compound 4 because this transition occurs during the excitation process (**Fig. 2.12**).

Computational identification of BODIHY dyes as FMRs. The compounds were then computationally assessed on their behavior as molecular rotors. The rotational barrier of the rotator (phenyl pendant) around the stator (BF_2-heterocyle) was determined for the dyes in a relaxed and excited state (**Table 2.1**).[23] The dihedral angle was increased in 10° increments to illustrate the full 360° rotation. Only for compound **2**, where methyl groups are present at the *ortho*- positions, was the energy barrier high enough to suggest impeded ring rotation due to steric interference. Upon inspection of the structure, it is unsurprising that any bulk at the *ortho*- position would drastically increase barrier to rotation (**Fig. 2.1**). All other screened dyes were determined to be molecular rotors as their low rotational barriers suggest unimpeded rotation in the gas phase. This gas phase rotation can be extrapolated to low viscosity solutions held at room temperature; the inclusion of COSMO continuum solvation model did not lead to substantial improvements in the calculated values, consequently the solvation model was omitted from subsequent calculations.[24]

The rate of rotation was determined by the Eyring equation to further quantify the molecular rotor behavior (**Eqn 2.2**).[25] Here, k_{rot} represents the rate of rotation, k_B represents the Boltzmann constant, T represents temperature, h represents Planck's constant, c° represents the standard concentration unit (1 mol/L), $\Delta G^{\ddagger}$ represents Gibbs energy of activation, and R represents the gas constant. The units for rate of rotation indicate a zeroth-order process. The value for $\Delta G^{\ddagger}$ was the difference between the final Gibbs free enthalpy of the transition state and the ground state.

$$k_{rot} = \frac{k_B T c^{\circ}}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}}$$

Equation 2.2

Compounds 1, 3, 4, 6, and 8 were found to have rotational rates on the GHz scale while rotational rate of compound 2 was on the order of Hz (**Table 2.8**). In a low viscosity solution at room temperature, compounds 1, 3, 4, 6, and 8 can be considered to behave as molecular rotors. This stark difference in rotational rate of compound 2 is unsurprising as the *ortho*-methyl groups are likely impeding free rotation of the pendant phenyl around the heterocycle moiety. For that reason, compound 2 is not a molecular rotor under the aforementioned conditions.

Table 2.8. TDDFT-computed rotational rates and barriers.

BODIHY	$\Delta E^{\ddagger}$ (kcal/mol)	$\Delta G^{\ddagger}$ (kcal/mol)	k_{rot} (M ns^{-1})
1	4.1	4.9	1.48
2	15.5	17.3	1.27×10^{-9}
3	5.0	5.1	1.14
4	5.1	5.6	0.50
6	4.4	4.9	1.49
8	5.2	6.6	0.09

SECTION 2.4. DISCUSSION

Suppression of Kasha's Rule. The experimentally verified computational modeling allowed for the better understanding of the AIE of the BODIHY dyes. Here we describe the quantum mechanical origin of the enhanced emission of the BODIHY dyes in viscous environments and in the solid state. The photophysical mechanism responsible for this AIE is called "Suppression of Kasha's Rule" (SOKR). [5]

The SOKR mechanism can best be explained by comparing the emissive properties of the compounds in two types of solvent environments: low viscosity and high viscosity. When the molecule is excited by an incident photon, low viscosity solvents would allow for unimpeded ring rotation into a conformation that favors IC to the S_1 excited state followed by non-radiative relaxation to S_0 (**Fig. 2.13**).

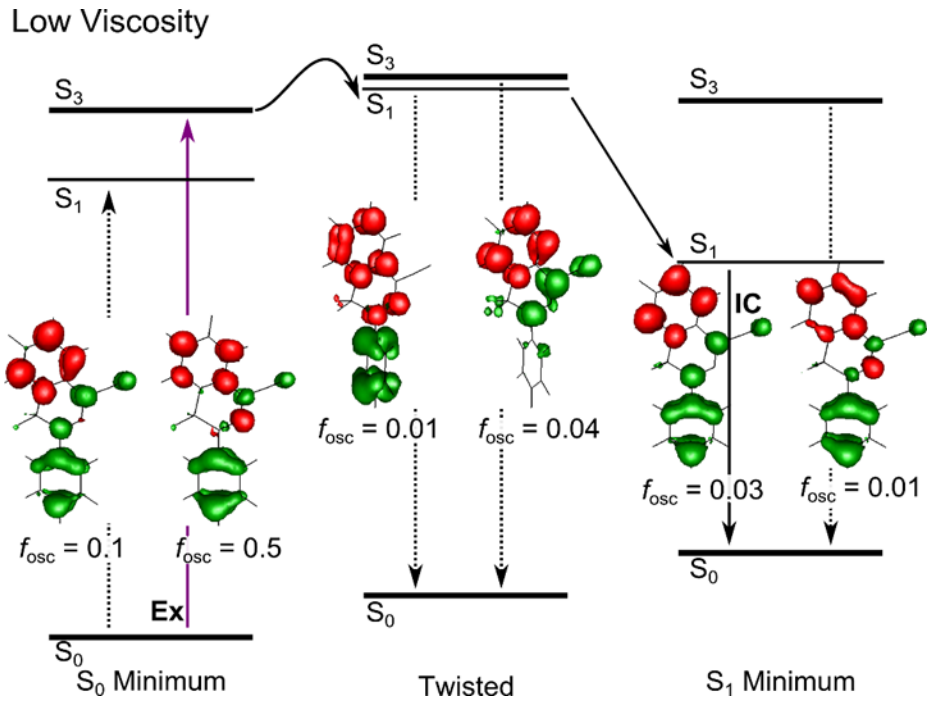


Figure 2.13. A simplified Jablonski diagram describing the SOKR mechanism resulting in emission enhancement of compound 1 in a low viscosity environment.

The resultant emission arises from the small percentage of molecules that radiatively relaxed from the bright state, $S_{>1}$, before IC to S_1; this explains the relatively low quantum yield of the dyes in low viscosity environments. When selectively exciting into the lower energy dark state, the weak emission at 577 nm occurs from the few molecules that relax radiatively from S_1.

When the molecule is excited by an incident photon, high viscosity solvents would prevent free ring rotation, thereby holding the FMR in the geometry from which it was excited. The hindrance of ring rotation resulting from highly viscous solutions will lead an increase in fluorescence intensity as depicted in **Fig. 2.14**.

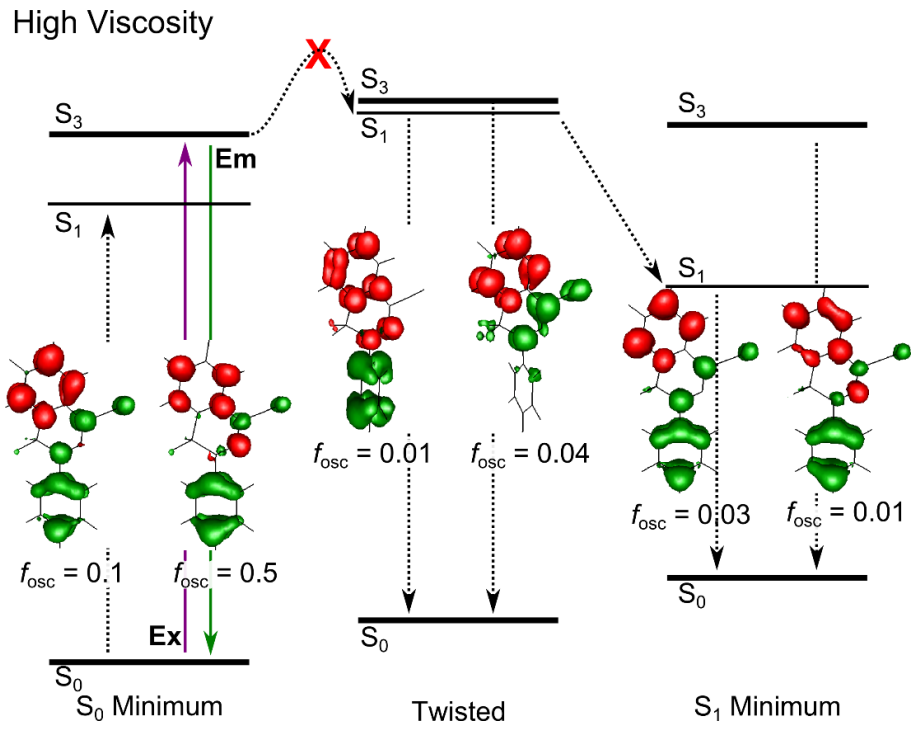


Figure 2.14. A simplified Jablonski diagram describing the SOKR mechanism resulting in emission enhancement of compound 1 in a high viscosity environment.

The hindered rotation encourages radiative relaxation from $S_{>1}$ as molecules will not reach the perpendicular conformation that allows for rapid IC to the dark state. The increased barrier of rotation introduced by more viscous solvents is paramount to understanding the enhancement of fluorescence intensity in these rotation-restrictive environments. Two factors contribute to barrier to rotation ($\Delta G^{\ddagger}$) for these FMRs: barrier due to solvent-solute interactions and barrier due to electronics. Because these FMRs have similar rotor bulk, the variable factor is the differences in electronics between the dyes.

More than one excited state identified through Abs and CD spectroscopies. To experimentally validate the proposed SOKR mechanism, a pair of spectroscopic techniques were used to better understand the viscosity dependent emission for compounds 1, 4, and 6. First, the existence of two states was confirmed through 77 K Abs spectroscopy and circular dichroism (CD) spectroscopy. The features found in these two spectra arise from the same set of electronic transitions. However, unique selection rules can alter the relative intensity of a feature in the Abs spectrum when compared to CD spectrum (**Eqn. 2.3** and **Eqn. 2.4**).

$$I(Abs) = \langle \Psi_i | \hat{\mu}_{el} | \Psi_f \rangle \qquad \textbf{Equation 2.3}$$

$$I(CD) = \langle \Psi_g | \hat{\mu}_{el} | \Psi_i \rangle \langle \Psi_g | \hat{\mu}_{mag} | \Psi_e \rangle \qquad \textbf{Equation 2.4}$$

The feature at 408 nm was present in both the Abs and CD spectra of compound 1 but these features had dissimilar bandshapes. A change in intensity of one (or more) transitions from the Abs spectrum to CD spectrum is responsible for the unique bandshapes. The two (or more) transitions comprising the feature at 408 nm must be very close in energy as the bandshapes are similar (**Fig. 2.15**).

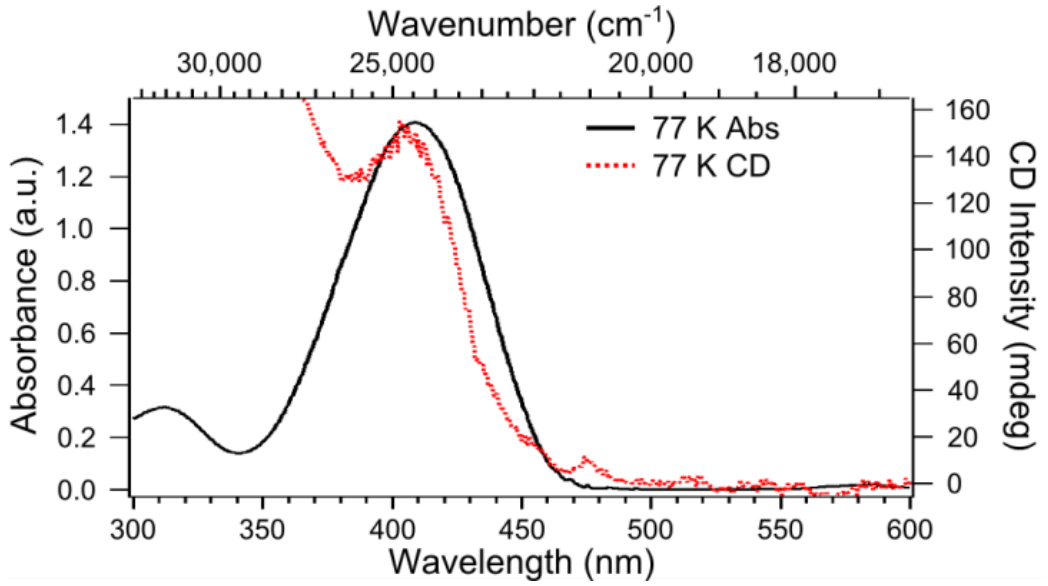


Figure 2.15. Overlain 77 K Abs (black) and CD (red) spectra of compound 1 (100 μM) in a frozen solution of 80% Gly and 20% EG (*v*/*v*%).

The rising baseline of the 77 K CD spectrum resulted from frost on the sample; this was corrected by a curve with a λ^{-4} dependence and is not indicative of a high-energy transition in the CD spectrum. The corresponding feature in the TDDFT-predicted Abs spectra is also comprised of two unique transitions (**Fig. 2.6**). Similar information was determined from the 77 K Abs and CD spectra collected for compound 6, which was selected as the ED parent compound (**Fig 2.16)**. The 77 K CD spectrum of the EW parent compound (compound 4) was featureless.[5] The featureless spectrum is likely due to the unique dominant electronic transition for this dye. Compound 4 is dominated by a BODIHY→NO_2 unlike the $\pi \rightarrow \pi^*$ transition of the other BODIHY dyes.

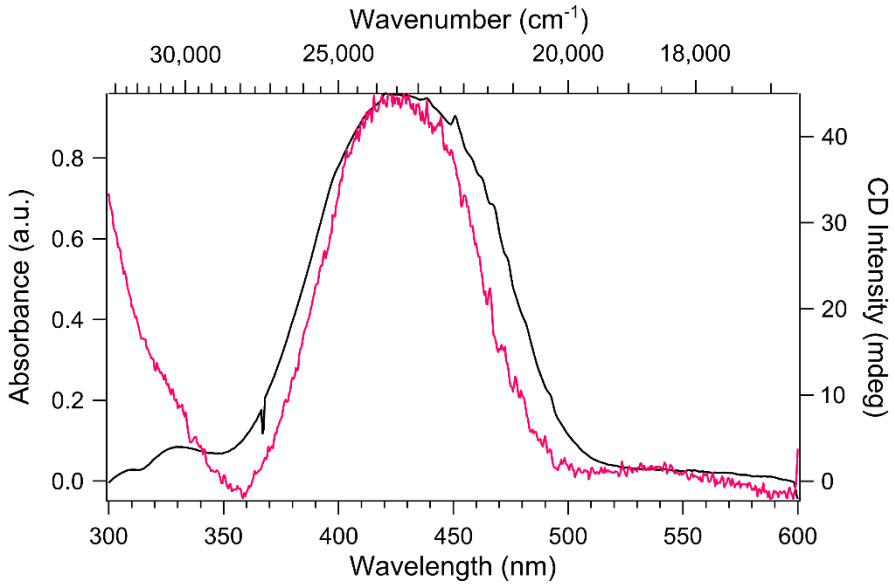


Figure 2.16. Overlain 77 K Abs (black) and CD (pink) spectra of compound 6 (70 μM) in a frozen solution of 80% Gly and 20% EG (*v*/*v*%).

In conclusion, comparing the Abs and CD spectra of compound 1 and 6 confirmed the existence of more than one relevant excited state. However, these two spectroscopic techniques cannot determine which excited state is responsible for the bright BODIHY emission.

Excitation-dependent emission spectroscopy identified two excited states. While the two electronic spectroscopies could only identify the existence of more than one state, fluorescence spectroscopy offers a more direct method to identify if the lower energy excited state, S_1, is a dark state as outlined in the SOKR mechanism. The major feature in the emission spectra (λ_{max} = 528 nm) was proposed to arise from the bright state, for compound 1 was predicted to be radiative relaxation from S_3 but for the sake of generality, $S_{>1}$ will be used to refer to the bright state. This single higher energy feature has been observed in multiple solvents of varied polarities and viscosities (albeit at different intensities due to SOKR).

To promote the detection of the dark state, a low viscosity solvent was chosen (dichloromethane, η= 0.4 cP) to increase the rate of IC from $S_{>1}$ to S_1 in order to increase the occurrence radiative relaxation from S_1. If the excitation wavelength is fixed to an energy lower than that responsible for the bright state, the energy should only be sufficient to excite into S_1. As the energy of the excitation wavelength was decreased, the disappearance of the bright state feature was followed by the appearance of the lower energy feature (λ = 577 nm) (**Fig. 2.17**).

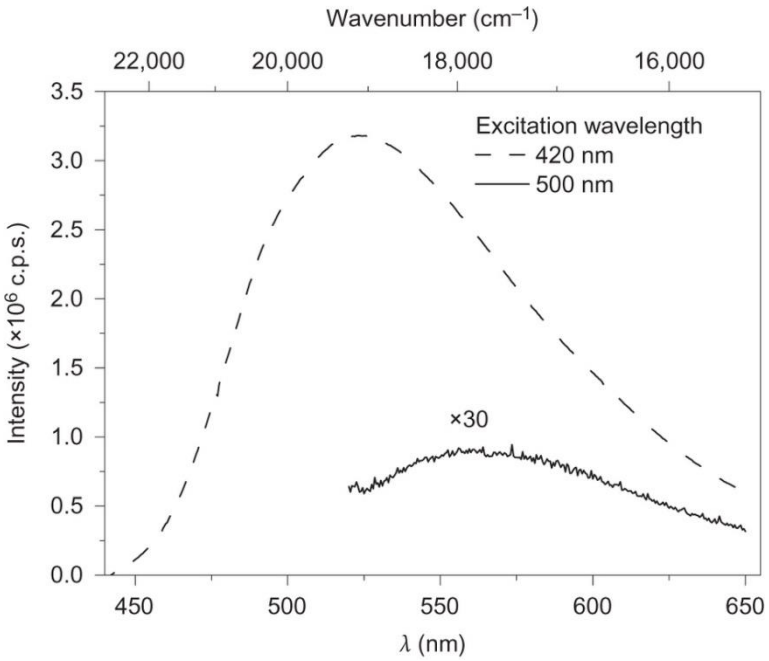


Figure 2.17. Excitation-dependent emission spectra of compound 1 in dichloromethane. The spectrum of dichloromethane was subtracted from the corresponding averaged spectra of the dye to remove the solvent's Raman scattering peaks and enhance the weak emission allowing it to be observed of the solvent interference.

This excitation wavelength dependence of the two features indicated that the features are unique and the lower energy emission feature must arise from the lower energy dark state. These data cement the proposal that Kasha's Rule, only having a bright state and a ground state, does not accurately describe the photophysical properties of this fluorophore family. Instead, a three-state system (ground, dark, and bright state) must be employed to explain the photophysical behaviors exhibited by the BODIHY dyes.

The SOKR mechanism is further supported by fluorescence lifetime data as well as quantum yield measurements.[5] This data was collected for compounds 1, 4, and 6 (**Table 2.9**). These experiments provided proof that both the dark state (S_1) and bright state ($S_{>1}$) were relevant to the photophysical mechanism behind BODIHY fluorescence. In solutions of decreasing viscosity, the rate of non-radiative relaxation significantly increased while the rate of radiative relaxation decreased. In low viscosity solutions, the lifetime (τ) and quantum-yield (Φ_{em}) measurements for compounds 1, 4, and 6 were lower indicating the non-radiative relaxation pathways were dominant (**Table 2.9**). For compound 4, the rate of non-radiative relaxation increased by an order of magnitude (this data was unattainable for compounds 1 and 6 because it was below the limit of detection for the instrument). The available data for compounds 1 and 6 (in the highest viscosity solvent cocktails) was consistent with the data of compound 4. The restriction of the rotation of the phenyl ring in a viscous environment leads to brighter emission because relaxation to the dark state is avoided.

Table 2.9. Fluorescence lifetime τ, quantum yield Φ_{em}, and the radiative k_r and nonradiative k_{nr} rate constants of compounds 1, 4, and 6 in solvents with different viscosities.

Compound 1				
η (cP)	τ (ns[a])	Φ_{em}[b]	k_r (ns^{-1})	k_{nr} (ns^{-1})
620.7	0.300	0.023	0.077	3.271
283.1	0.177	0.01	0.056	5.544
124.0	0.103	-[c]		
80.8	0.076	-[c]		
52.0	0.070	-[c]		
20.8	0.049	-[c]		
Compound 4				
η (cP)	τ (ns[a])	Φ_{em}[b]	k_r (ns^{-1})	k_{nr} (ns^{-1})
620.7	0.572	0.110	0.192	1.556
283.1	0.348	0.061	0.175	2.698
124.0	0.214	0.036	0.168	4.505
80.8	0.163	0.025	0.153	5.982
52.0	0.141	0.019	0.135	6.957
20.8	0.091	0.013	0.143	10.846
Compound 6				
η (cP)	τ (ns[a])	Φ_{em}[b]	k_r (ns^{-1})	k_{nr} (ns^{-1})
620.7	0.515	0.035	0.068	1.875
283.1	0.344	0.023	0.067	2.846
124.0	0.228	0.01	0.044	4.356
80.8	0.190	-[c]		
52.0	0.151	-[c]		
20.8	0.111	-[c]		

[a] 1.2% error.
[b] 1.5% error.
[c] below detection limit.

SECTION 2.5. CONCLUSION

The data presented here support a novel mechanism to describe the AIE of the family of BODIHY dyes: Suppression of Kasha's Rule. It has been proposed that other anomalous fluorophores (or phosphors) that have viscosity sensitive emission profiles and AIE may also be described by the SOKR mechanism. Fluorophores must display AIE, anomalous emission, and a barrier to rotation in the excited state. A two-fold approach has been taken to identify molecules that could be SOKR fluorophores: (1) BODIHY derivatives and (2) previously documented anomalous fluorophores. These two approaches will be discussed in later chapters of this dissertation.

SECTION 2.6. REFERENCES

1. Yang, Y.; Su, X.; Carroll, C. N.; Aprahamian, I., *Chem. Sci.* **2012,** *3*, 610-613.

2. Haidekker, M. A.; Theodorakis, E. A., *J. Biol. Eng.* **2010,** *4*, 11.

3. Hong, Y.; Lam, J. W.; Tang, B. Z., *Chem. Commun.* **2009**, 4332-4353.

4. Birks, J. B., *Photophysics of aromatic molecules*. Wiley-Interscience: London, New York, 1970; p xiii, 704 pages.

5. Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., *Nat. Chem.* **2017,** *9*, 83-87.

6. Haidekker, M. A.; Brady, T. P.; Lichlyter, D.; Theodorakis, E. A., *Bioorg. Chem.* **2005,** *33*, 415-425.

7. Förster, T.; Kasper, K., *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics* **1954,** *1*, 275-277.

8. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. 2nd ed.; Plenum Publishers: New York, NY, 1999.

9. Luo, J.; Xie, Z.; Lam, J. W.; Cheng, L.; Chen, H.; Qiu, C.; Kwok, H. S.; Zhan, X.; Liu, Y.; Zhu, D.; Tang, B. Z., *Chem. Commun.* **2001**, 1740-1741.

10. Mei, J.; Leung, N. L.; Kwok, R. T.; Lam, J. W.; Tang, B. Z., *Chem. Rev.* **2015,** *115*, 11718-11940.

11. Born, M.; Oppenheimer, R., *Ann. Phys.* **1927,** *389*, 457-484.

12. Haidekker, M. A.; Theodorakis, E. A., *J. Mater. Chem. C* **2016,** *4*, 2707-2718.

13. Allen, B. D.; Benniston, A. C.; Harriman, A.; Rostron, S. A.; Yu, C., *Phys. Chem. Chem. Phys.* **2005,** *7*, 3035-3040.

14. Neese, F., *Wiley Interdiscip Rev: Comput Mol Sci* **2012,** *2*, 73-78.

15. Perdew, J. P.; Burke, K.; Ernzerhof, M., *Phys. Rev. Lett.* **1996,** *77*, 3865-3868.

16. Schäfer, A.; Horn, H.; Ahlrichs, R., *J. Chem. Phys.* **1992,** *97*, 2571-2577.

17. Zhou, Y.; Xiao, Y.; Chi, S.; Qian, X., *Org. Lett.* **2008,** *10*, 633-636.

18. Kasha, M., *Disc. Faraday Soc.* **1950,** *9*, 14-19.

19. Adamo, C.; Scuseria, G. E.; Barone, V., *J. Chem. Phys.* **1999,** *111*, 2889-2899.

20. Chai, J. D.; Head-Gordon, M., *J. Chem. Phys.* **2008,** *128*, 1-16.

21. Itoh, T., *Chem. Rev.* **2012,** *112*, 4541-4568.

22. Munkholm, C.; Parkinson, D. R.; Walt, D. R., *J. Am. Chem. Soc.* **1990,** *112*, 2608-2612.

23. Kottas, G. S.; Clarke, L. I.; Horinek, D.; Michl, J., *Chem. Rev.* **2005,** *105*, 1281-1376.

24. Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F., *J. Phys. Chem. A* **2006,** *110*, 2235-2245.

25. Eyring, H., *J. Chem. Phys.* **1937,** *3*, 107-115.

CHAPTER 3. IDENTIFYING NEW SOKR MOLECULES: DIHALIDE HYDRAZONE DERIVATIVES

SECTION 3.1. INTRODUCTION

Expanding the molecules that fall under the Suppression of Kasha's Rule (SOKR) mechanism. SOKR molecules are examples of aggregation induced emission (AIE) fluorophores. A SOKR fluorophore has a barrier to rotation in the emissive state that is sensitive to the solvent environment, specifically, viscosity. To date, the boron difluorohydrazone (BODIHY) dyes are the only members of the SOKR molecule family. It has been proposed that the AIE of other fluorophores can be explained by the SOKR mechanism. Thus, new hydrazone-based dyes were the initial target to identify new SOKR fluorophores. New derivatives of BODIHY were screened computationally first before highlighting the most attractive synthetic targets for our collaborators (Professor Ivan Aprahamian and coworkers, Dartmouth College, Department of Chemistry) (**Fig. 3.1**)

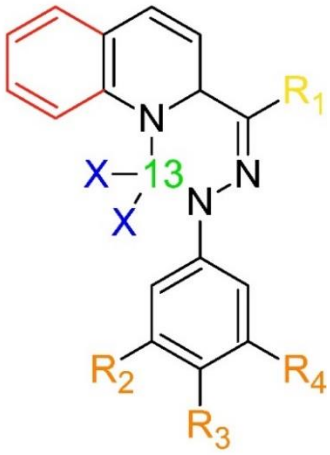


Figure 3.1 General structure for new hydrazone derivatives. The heterocycle can be expanded to a naphthyl ring (red). Different electron withdrawing or donating groups can be substituted at the R_1 position (yellow). Group 13 elements (green: Al, Ga, In, Tl) can replace boron. Heavier halides (blue X= Cl, Br, or I) can replace fluorine. The pendant phenyl can be substituted at any position other than the *ortho*- position (orange: R_2-R_4).

There are many sites that can be changed through the tunable synthesis for the hydrazone dyes: heterocycle group, R_1 group, the halogens, substituents on the phenyl pendant, and the group 13 element. In previous works, the Aprahamian group synthesized the boron difluorohydrazone dyes with different aryl groups expanding the heterocycle phenyl with naphthyl rings.[1] Increasing the π-system of the hydrazone dye led to rigidity in the solid state; the boron-containing heterocycle was nearly perpendicular to the pendant phenyl ring in the solid state (CCDC: 841446). The quantum yield (Φ_F) of this naphthyl derivative in solution and the solid state was very low; therefore, BODIHY derivatives with expanded π-systems will not explored in this chapter.[1] Changing the R_1 group greatly influences the emission efficiency in the solid state. The hydrazone with a cyano group for

R_1 has the greatest efficiency when compared to methyl, acetyl, or methoxycarbonyl group at the R_1 position.[1]

Three handles will be investigated computationally to build a library of synthetic targets for our collaborators: substituents on the phenyl pendant, the halogens, and the group 13 element (**Fig 3.1**). The effect of the substituents on the phenyl ring influenced the photophysical properties of the BODIHY dyes; this handle allows for a straight-forward synthetic approach to tune to color of the BODIHY emission.[2] Substituents will be added to the *para*-, *meta*-, or both *meta*- positions. Based on computational and experimental data, substitutions at the *ortho*-position were not explored because the barrier to rotation is too high and these dyes would not behave like molecular rotors.[2] Heavier halogens can be substituted and may induce intersystem crossing (ISC) due to increased spin-orbit coupling thus creating new phosphorescent hydrazone derivatives.[3] The last handle to be changed would be the boron group; heavier group 13 elements could also lead to phosphorescent hydrazone dyes.[3]

Motivation for new BODIHY derivatives. One motivation for computationally exploring these compounds was to identify a new synthetic target that would be as bright as *para*-NO_2 BODIHY (compound 4). Compound 4 was shown to have high viscosity sensitivity, a large Stokes shift, and was the brightest compound synthesized to date.[2] However in the solid state, compound 4 forms H-aggregates and fluorescence is quenched.[2] Quenching in the solid state is a poor attribute for a fluorophore; this severely limits the practical application in endeavors such as the development of organic light emitting diodes (OLEDs) or the use of thin films.[4] Ideally, a new BODIHY derivative will have similar

photophysical properties as compound 4 but retain fluorescence in the solid state. New substituents, specifically the addition of new substituents at the *meta*- or *para*- positions, should alter the manner of packing. Larger group 13 elements will lead to puckering of the heterocycle, which will affect packing.

Hydrazone derivatives will be investigated for their molecular rotor behavior, predicted brightness, as well as potential behavior as a phosphorescent dye (**Fig 3.2**). As targets are synthesized by our collaborators, this roadmap will be updated to allow for more accurate predictions. A series of fluorophore and phosphor targets were screened computationally and the most promising boron and aluminum hydrazones will be highlighted in this chapter.

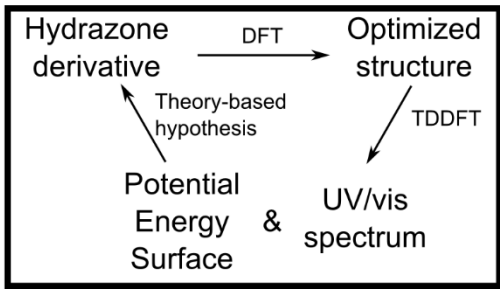


Figure 3.2. Roadmap for identifying potential synthetic targets derived from the hydrazone skeleton.

SECTION 3.2. METHODS

Computational approach to identifying attractive synthetic targets. These new targets were modeled by density functional theory (DFT) and time-dependent DFT (TDDFT) as previously described (PBE functional, TZVP basis set, and tight SCF convergence).[2] For the BODIHY dyes, there was minimal influence on the predicted parameters when the COSMO continuum solvation model so it was omitted for the hydrazone derivative calculations. The crystal structure of unsubstituted BODIHY was used as the initial structures for building these compounds on ArgusLab (Planaria Software).[1] A geometry optimization calculation was completed to identify the minimum energy geometry of the BODIHY derivative. The targeted boron and aluminum hydrazone derivatives are listed in **Table 3.1**.

Table 3.1. Hydrazone derivatives modeled by DFT and TDDFT.

Boron	Aluminum
BF_2-hydrazones	AlF_2-hydrazones
No substituents	No substituents
-NO_2 (3; 3,5)	-NO_2 (3; 3,5; 4)
-CN (3; 3,5)	$AlCl_2$-hydrazones
-F (3; 3,5; 4)	No substituents
-Br (3; 3,5; 4)	-NO_2 (3; 3,5; 4)
-Cl (3; 3,5; 4)	$AlBr_2$-hydrazones
-iPr (3; 3,5; 4)	No substituents
-tBu (3; 3,5; 4)	-NO_2 (3; 3,5; 4)
BBr_2-hydrazones	$AlCl_2$-hydrazones
No substituents	No substituents
-NO_2 (3; 3,5; 4)	-NO_2 (3; 3,5; 4)
BI_2-hydrazones	
No substituents	
-NO_2 (3; 3,5; 4)	
3= *meta*- substituent on the phenyl 3,5= both *meta*- substituents of the phenyl 4= *para*- substituent on the phenyl	

Behavior as a molecular rotor. Like the BODIHY dyes, the DFT models of hydrazone derivatives as a function of rotor angle were comprised of a series of constrained geometry optimizations. These calculations were run in the gas phase and were interpreted to represent the environment of a low viscosity solvent at room temperature. The dihedral angle was increased in 10° increments to mimic the 360° rotation of the hydrazone. The dihedral angle of the molecular rotor was held constant while the remaining bond lengths

and angles were allowed to relax (**Fig. 3.3**). The ground state barrier to rotation was the difference in energy between the highest energy conformation and the lowest energy conformation.

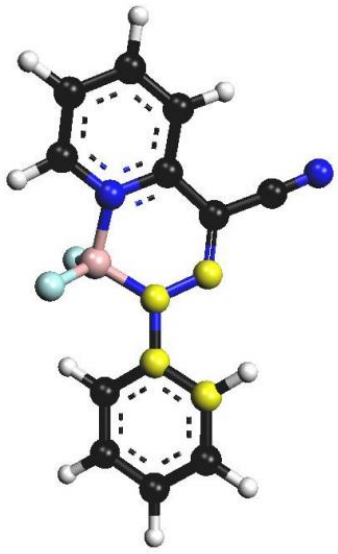


Figure 3.3. Structure of a generic hydrazone used for geometry optimizations generated in ArgusLab (Planaria Software). Black spheres represent carbon, blue spheres represent nitrogen, pink spheres represent group 13 element, light blue spheres represent halogens and white spheres represent hydrogen atoms. Substitutions may occur on the hydrazone as previously described. The highlighted N–N–C–C (yellow atoms) dihedral angle was varied in 10° increments.

Brightness and color-tuning of predicted molecular rotors. The preliminary investigation of brightness and color of the modeled dyes was based on the TDDFT-predicted UV-vis absorption (Abs) spectra of each hydrazone derivative. After the energy of the structure was minimized, the electronic structure of the dye was modeled by PBE TDDFT. The Abs spectra of the dyes were simulated based upon the TDDFT data by convoluting Gaussian-shaped bands with full width at half maximum bandwidths of 4000 cm^{-1} using the orca_mapspc utility program. This bandwidth was derived from the experimental bandwidth of boron difluorohydrazone in ethylene glycol.[2] The color of each *para*-substituted BODIHY dye was derived from the energy of the lowest energy feature in the predicted Abs spectrum. The brightness of each molecular rotor was ranked based on the predicted molar extinction coefficients of the main transition of the TDDFT-predicted Abs spectrum (**Eqn. 3.1**).

Barrier to rotation in the excited state. The first twenty excited states for each of the 36 different geometries were predicted by TDDFT. The excited state with the highest oscillator strength was selected as the emissive excited state. The excited state barrier to rotation was the difference in energy between the highest energy conformation and the lowest energy conformation. These calculations were run in the gas phase and were interpreted to represent the environment of a low viscosity solvent at room temperature.

SECTION 3.3. RESULTS

Are the potential targets molecular rotors? To be categorized as a potential SOKR luminogen, the hydrazones must behave like molecular rotors in solution. Compounds with computed ground state barriers greater than 10 kcal mol^{-1} were not expected to behave as molecular rotors in solution at room temperature. Only molecules that are predicted to rotate freely in solution would be investigated further. As no solvation model was implemented, these gas phase calculations mimic the behavior of the hydrazone in a low viscosity solvent at room temperature. For reference, the DFT-predicted barrier to rotation for boron difluorohydrazone without any substitutions was 4.1 kcal mol^{-1}.[2] A series of EW groups (-NO_2, -CO_2CH_3, -CO_2H), ED (-CN, -iPr, -tBu, -OH) groups as well as halogens (inductive EWGs) were added to the *para*-, *meta*-, or both *meta*- positions of the phenyl ring (**Table 3.2**). Two hydrazones, *para*-cyano BODIHY and *para*-NO_2 BODIHY, have been previously characterized.[2]

Table 3.2. Predicted ground state energy barriers for a series of boron difluorohydrazone dyes.

BF_2-hydrazones			
para-substituent	R_2	R_2 and R_4	R_3
CN	4.4	4.6	*para*-CN BODIHY
iPr	4.0	4.0	4.2
tBu	4.1	4.0	4.2
NO_2	4.5	4.7	*para*-NO_2 BODIHY
F	4.3	4.4	4.1
Cl	4.2	4.2	4.4
Br	4.2	4.2	4.4
OH	-	-	4.4
CO_2H	-	-	4.8
CO_2CH_3	-	-	4.7

The heavy-atom effect can increase the rate of ISC leading to emission from the triplet state.[3] Substituting fluorine with heavier halogens may be a pathway to synthesize phosphorescent dyes. The replacement of boron with aluminum also provided a new set of fluorescent and potentially phosphorescent molecular rotors. The new halogens or group 13 elements did not cause significant puckering of the hydrazone dye and were all predicted to behave as molecular rotors (**Table 3.3**). The optimized structures were viewed in ArgusLab (Planaria software).

Table 3.3. Predicted ground state energy barriers for select boron- and aluminum-based hydrazones.

Hydrazone	G.S. Barrier (kcal/mol)
BCl_2-hydrazone	2.5
4-NO_2 BCl_2-hydrazone	3.2
BCl_2-hydrazone	2.8
4-NO_2 BCl_2-hydrazone	3.1
BI_2-hydrazone	4.2
4-NO_2 BI_2-hydrazone	4.2
AlF_2-hydrazone	4.9
4-NO_2 AlF_2-hydrazone	6.0
$AlCl_2$-hydrazone	3.9
4-NO_2 $AlCl_2$-hydrazone	4.9

Substitution at the *para*-position of the pendant phenyl of BODIHY allows for color tuning. One motivation to computationally investigate more hydrazone derivatives is to expand the color palate of the emission of the BODIHY family. Also, the expansion of the family of synthesized dyes to include these proposed hydrazones will gradually improve our prediction model. The PBE functional correctly predicted the energy shift of the Abs maxima upon addition of electron withdrawing and donating groups at the para position for BODIHY.[2] The new BODIHY dyes were predicted to have molar extinction coefficients comparable to the synthesized BODIHYs (**Table 3.4**).

Table 3.4. TDDFT predicted Abs maxima and molar extinction coefficients for hydrazone derivatives with *para*-substituted phenyl rings.

BODIHY	λmax (nm)	ε (M^{-1} cm^{-1})
BF_2-hydrazone[a]	403	28000
-CO_2CH_3 [b]	402	28900
-CO_2H [b]	405	33800
-NO_2 [a]	414	40800
-CN [a]	414	28600
-F [b]	417	26700
-Br [b]	459	22400
-Cl [b]	436	25100
-OH [b]	453	21800
-iPr [b]	424	29700
-tBu [b]	425	29200
-OCH_3 [a]	469	26200
-$N(CH_3)_2$ [a]	530	22600

a. Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., *Nat Chem* 2017, *9*, 83-87.
b. This chapter

TDDFT-predicted brightness of fluorophores. A series of boron and aluminum hydrazone dyes were investigated for the brightness of their emission through TDDFT calculations. The predicted molar extinction coefficient (ε) of most intense Abs feature is related to the brightness (α) of the fluorophore (**Eqn. 3.1**).[5]

$$\alpha \approx \varepsilon \cdot \Phi_F \qquad \textbf{Equation 3.1}$$

The introduction of heavier halogens as well as group 13 elements BODIHY derivative was predicted to alter the energy and intensity of the main Abs feature (**Table 3.5**). The heavy atom effect will increase the likelihood of intersystem crossing to the triplet state.[3] Some of these hydrazone are likely to be phosphorescent.

Table 3.5. TDDFT-predicted Abs spectra parameters of BODIHY and ALDIHY dyes.

Hydrazone derivative	λmax (nm)	ε (M^{-1} cm^{-1})
BCl_2-hydrazone	420	15200
4-NO_2 BCl_2-hydrazone	413	24600
BBr_2-hydrazone	409	13700
4-NO_2 BBr_2-hydrazone	431	21700
BI_2-hydrazone	402	10800
4-NO_2 BI_2-hydrazone	486	13900
AlF_2-hydrazone	407	31900
4-NO_2 AlF_2-hydrazone	413	45100
$AlCl_2$-hydrazone	416	27000
4-NO_2 $AlCl_2$-hydrazone	416	37700
$AlBr_2$-hydrazone	418	23700
4-NO_2 $AlBr_2$-hydrazone	407	35800
AlI_2-hydrazone	417	23900
4-NO_2 AlI_2-hydrazone	420	34000

Barrier to rotation in the emissive excited state. The rotation barrier in the excited state was only calculated for the most promising hydrazone dyes. The most promising dye, *para*-NO_2 aluminum difluorohydrazone, had the largest molar extinction coefficient (**Table 3.5**). This dye was predicted to be brighter than compound 4 which the brightest synthesized to date. To better understand the boron-to-aluminum substitution, the aluminum difluorohydrazone was also computationally investigated (**Table 3.6**).

Table 3.6. Barrier to rotation in G.S. and E.S. for select hydrazones.

BODIHY	G.S. Barrier (kcal/mol)	E.S. Barrier (kcal/mol)
BF_2-hydrazone	4.1	7.1
para-NO_2 BF_2-hydrazone	5.1	7.7
AlF_2-hydrazone	4.9	2.6
para-NO_2 AlF_2-hydrazone	6.0	8.3

SECTION 3.4. DISCUSSION

All dyes were predicted to behave as molecular rotors. Free rotation in a room-temperature solution is required for these dyes to be classified as molecular rotors (**Table 3.2** and **3.3**). By avoiding substitutions at the *ortho*- positions of the phenyl pendant, the phenyl can freely rotate around the heterocycle. In the future, the inclusion of larger substituents at the *meta*-position on the phenyl ring or larger group 13 elements may increase the ground state rotation barrier to above 10 kcal mol^{-1}. Significant structural changes may occur with the larger atoms in the heterocycle.

Tuning color by modifying *para*-position on phenyl ring. In accordance with the previously published data, the addition of EW groups led to hypsochromic shifts while ED groups had bathochromic shifts compared to the unsubstituted BODIHY (**Table 3.4**). The halogen groups can act as either EW or ED groups. The inclusion of non-bonding lone pair electrons of the halogens can donate into the π-system of the phenyl ring. For example, *para*-bromo and *para*-chloro BODIHY were predicted to red-shift to 459 and 436 nm respectively. Fluorine is more electronegative than bromine and chlorine and does not follow this trend; the heavier halogens are likely acting as electron donating groups.

The Abs maxima of the published BODIHY dyes fall between 20,660 and 26,950 cm^{-1} (484 and 371 nm).[2] The new dyes discussed in this chapter have predicted Abs maxima fall between 21,320 and 24,880 cm^{-1} (**Table 3.4**). These new dyes will expand the blue color palate of the BODIHY family. While no *para*-substituted BODIHY dye was predicted to have a larger molar extinction coefficient than *para*-NO_2 BODIHY, a large quantum yield could raise the brightness of these dyes to be greater than that of *para*-NO_2

BODIHY (**Table 3.4**). Computationally screening the family of dyes with more ED and EW groups will lead to more focused dye synthesis. The prediction method will be improved by comparing the experimental data of the dyes synthesized by our collaborators to the DFT- and TDDFT-predicted parameters.

Identifying a new fluorophore to replace *para*-NO_2 boron difluorohydrazone. Due to packing in the solid-state, the emission of compound 4 (brightest fluorophore in solution) was quenched.[2] Overall, the aluminum based hydrazones were predicted to high a higher molar extinction coefficient. The addition of the *para*-NO_2 group onto the phenyl ring consistently increases molar extinction coefficient and lowers the energy of the main feature (**Table 3.5**).

One fluorophore was predicted to be brighter than *para*-NO_2 BODIHY: *para*-NO_2 aluminum difluorohydrazone (**Table 3.5**). The parent compound of this dye, unsubstituted aluminum difluorohydrazone, was also investigated to assess the influence of an aluminum group replacing the boron. The only new BODIHY derivative to have a predicted molar absorptivity greater than compound 1 was *para*-CO_2H BODIHY. This dye, while not predicted to be brighter than compound 4, is still an attractive synthetic target.

Deter H-aggregation with a larger group 13 element. The replacement of boron for aluminum is a promising route as these aluminum derivatives are expected to pack differently in the solid state. To understand packing changes, the crystal structures of BODIHY and *para*-NO_2 BODIHY were compared to the geometry optimized structures of their aluminum derivatives; the crystal structures were viewed with the Mercury freeware and the computed structures were viewed in ArgusLab. The lowest geometry predicted for

heterocycle of AlF_2-hydrazone is more planar than the heterocycle in the BF_2-hydrazone crystal structure. The quantum yield of the dye should be enhanced in the solid state if the dye deviates from a perpendicular geometry.[1] Because the aluminum-derivative is more planar, this may result in a different packing conformation as compared to BODIHY. The predicted structure of *para*-NO_2 AlF_2-hydrazone is more puckered (in the heterocycle) and the phenyl ring is less in-plane with the heterocycle when compared to the crystal structure of *para*-NO_2 BF_2-hydrazone. It is likely that the less planar nature of the *para*-NO_2 AlF_2-hydrazone monomer will lead to a different solid-state packing that may not form deleterious H-aggregates as seen in compound 4.

Viscosity sensitivity related to barrier to rotation in excited state of aluminum difluorohydrazone dyes. In accordance with the SOKR mechanism, there is a strong correlation between the excited state energy barriers and viscosity sensitivity.[2] Aluminum difluorohydrazone has very low excited state barrier to rotation (**Table 3.5**). This barrier to rotation is lower than the lowest excited state barrier for the BODIHY dye (compound 8); the emission of AlF_2-hydrazone should not exhibit significant viscosity sensitivity.[2] Like compound 8, the aluminum difluorohydrazone dye is predicted to have a low brightness value. For AlF_2-hydrazone, it is predicted that the third excited state will be the brightest state with the highest oscillator strength (**Fig. 3.4**).

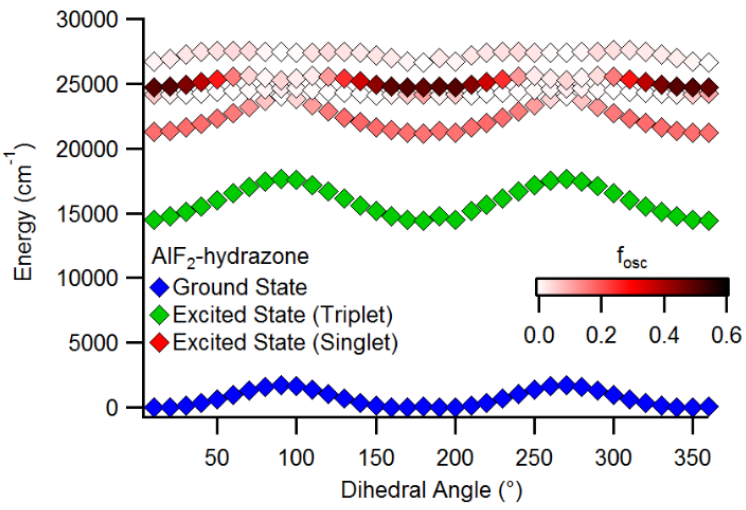


Figure 3.4. The potential energy surface PBE TDDFT-predicted ground (blue diamonds), singlet excited states (shaded red diamonds), and triplet excited state (green diamonds) of aluminum difluorohydrazone.

Anomalous emission is predicted to occur from S_3 for aluminum difluorohydrazone which breaks Kasha's Rule.[6,7] Relaxation from S_1 is predicted to be predominantly non-radiative as suggested by the weak oscillator strength. As with the BODIHY dyes, it can be concluded that enhanced emission is directly related to the conformation of the molecule based on this potential energy surface. The DFT- and TDDFT-predicted photophysical properties of aluminum difluorohydrazone support that this dye is a SOKR fluorophore. Upon successful synthesis of the hydrazone by our collaborators, this will be tested experimentally.

There is an energetically-accessible triplet state, T_1, from which phosphorescence is possible (**Fig. 3.4**). The triplet state may be accessible as per the heavy atom effect.[3] It is possible that the heavier group 13 element can increase the SOC to increase the rate of

ISC. This must be determined experimentally. The barrier to rotation in the lowest energy triplet excited state is 9.4 kcal mol^{-1}. By freezing the aluminum difluorohydrazone in a matrix, phosphorescence from T_1 should be detectable.[7] However, an accessible triplet state does not guarantee phosphorescence. The rate of intersystem crossing to allow relaxation from T_1 may not be competitive compared to radiative relaxation from the singlet excited state.

The DFT- and TDDFT-predicted photophysical properties of *para*-NO_2 aluminum difluorohydrazone differ from its unsubstituted precursor. For the *para*-NO_2 AlF_2-hydrazone species, there are two bright states (S_2 and S_3) that could predominate at certain conformations (**Fig. 3.5**).

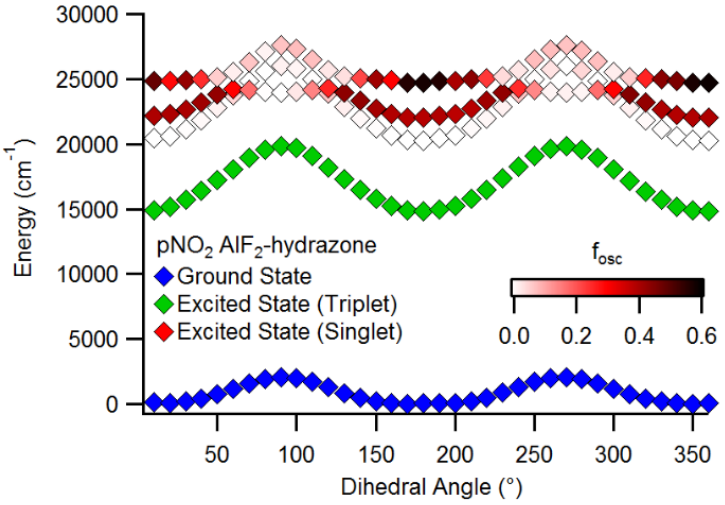


Figure 3.5. The potential energy surface PBE TDDFT-predicted ground (blue diamonds), singlet excited states (shaded red diamonds), triplet excited state (green diamonds) of *para*-NO_2 aluminum difluorohydrazone.

The relatively high barrier of rotation of the excited state of *para*-NO_2 AlF_2-hydrazone is a promising attribute implying high viscosity sensitivity. With an E.S. barrier (S_3) greater than that of compound 4, the emission enhancement of *para*-NO_2 AlF_2-hydrazone should have a high dependence on solvent viscosity (**Table 3.6**). The E.S. barrier for the second excited state is 11.5 kcal mol^{-1}; unimpeded rotation is unlikely to occur at this excited state in solution at room temperature.

SECTION 3.5. CONCLUSIONS AND FUTURE WORK

Tunable handles influence photophysical properties of the BODIHY dyes. The hydrazone derivatives were assessed on their ground state and excited state barriers of rotation, potential energy surfaces, as well as TDDFT-predicted Abs spectra. All screened targets were predicted to behave as molecular rotors in solution at room temperature (G.S. barriers were less than 10 kcal mol^{-1}). Predicting the photophysical properties of a series of dyes allows for a more efficient approach for selecting which fluorophore to synthesize. This is not only useful for nominating potential SOKR fluorophores, but also for general luminogens. This chapter has described an approach that:

(1) Determined if a hydrazone derivative will behave like a molecular rotor
(2) Related TDDFT-predicted molar excitation coefficient to the brightness of the fluorophore
(3) Showed that color-tuning can be achieved by altering the *para*-substituent on the phenyl ring
(4) Predicted a barrier to rotation in the emissive state which is correlated to viscosity sensitivity
(5) For heavier halogen BODIHY and ALDIHY compounds, phosphorescence is proposed by the predicted triplet states

Dyes can be developed with tailored photophysical properties that are predicted with electronic structure calculations. As more dyes are synthesized, a more accurate computational model will be developed.

Most attractive hydrazone derivative. The most attractive BODIHY was the *para*-CO_2H BF_2-hydrazone; this was the only proposed hydrazone in this chapter with a larger molar absorptivity than unsubstituted BODIHY. As a concluding remark pertaining to *para*-NO_2 AlF_2-hydrazone, this compound should be brighter than compound 4 with a more viscosity dependent emission. We also have reason to believe, due to the differences in the geometry as compared to compound 4, that this molecular rotor may not form H-aggregates in the solid state and may be fluorescent in the solid state.

Creating SOKR phosphors through heavy atom effect. It has been proposed that synthetic targets where the boron group has been substituted with other group 13 elements may behave as viscosity sensitive phosphors. The molecular rotor behavior may be interrupted by the larger atoms; the planarity of the fluorophore or phosphor may be disrupted. The heavy atom effect leads to an increase in spin-orbit coupling thus leading to intersystem crossing from the singlet excited state to a triple excited state.[3] Phosphors have a longer lifetime which could be a useful aspect as a microscopic viscosity sensors or other materials application. Phosphorescent materials also have the advantage of being more resistant to photobleaching.[8] Derivatives with heavier atoms will be made for the halogens as well as the other Group 13 elements and will be computationally screened for the aforementioned photophysical properties (**Table 3.7**).

Table 3.7. Potential hydrazone phosphors with heavy group 13 elements.

Gallium	Indium	Thallium
GaF_2-hydrazones	InF_2-hydrazones	TlF_2-hydrazones
No substituents	No substituents	No substituents
4-NO_2	4-NO_2	4-NO_2
$GaCl_2$-hydrazones	$InCl_2$-hydrazones	$TlCl_2$-hydrazones
No substituents	No substituents	No substituents
4-NO_2	4-NO_2	4-NO_2
$GaBr_2$-hydrazones	$InBr_2$-hydrazones	$TlBr_2$-hydrazones
No substituents	No substituents	No substituents
4-NO_2	4-NO_2	4-NO_2
GaI_2-hydrazones	InI_2-hydrazones	TlI_2-hydrazones
No substituents	No substituents	No substituents
4-NO_2	4-NO_2	4-NO_2

For both aluminum difluorohydrazone and *para*-NO_2 aluminum difluorohydrazone, a triplet state is present. With this information, phosphorescence can be proposed but experimental evidence is required to verify that claim. It is expected that the addition of heavier halides and group 13 elements would lead to a decrease in fluorescence. This decreased fluorescence is rooted in the increased rate of ISC and rate of phosphorescence.[9]

SECTION 3.6. CHAPTER 3 REFERENCES

1. Yang, Y.; Su, X.; Carroll, C. N.; Aprahamian, I., *Chem. Sci.* **2012,** *3*, 610-613.

2. Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., *Nat. Chem.* **2017,** *9*, 83-87.

3. McGlynn, S. P.; Daigre, J.; Smith, F. J., *J. Chem. Phys.* **1963,** *39*, 675.

4. Hong, Y.; Lam, J. W.; Tang, B. Z., *Chem. Commun.* **2009**, 4332-4353.

5. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. 2nd ed.; Plenum Publishers: New York, NY, 1999.

6. Kasha, M., *Disc. Faraday Soc.* **1950,** *9*, 14-19.

7. Itoh, T., *Chem. Rev.* **2012,** *112*, 4541-4568.

8. Wang, H.; Zhao, E. G.; Lam, J. W. Y.; Tang, B. Z., *Materials Today* **2015,** *18*, 365-377.

9. Koziar, J. C.; Cowan, D. O., *Acc. Chem. Res.* **1978,** *11*, 334-341.

CHAPTER 4. DO OTHER ANOMALOUS FLUOROPHORES EXHIBIT AGGREGATION INDUCED EMISSION?

SECTION 4.1. INTRODUCTION

Emission from a higher lying excited state has been documented for fluorophores and phosphors. Kasha's Rule dictates that upon excitation, any radiative relaxation will occur from the first excited state (S_1 for fluorophores and T_1 for phosphors).[1] While Kasha's Rule does accurately describe the photophysical behavior of many fluorophores, exceptions to Kasha's Rule have been reported. In 1955, azulene (Az) was the first documented fluorophore to display emission from a higher energy excited state (S_2).[2] Many publications since the 1950s have supported the claim that the emission from many fluorophores and phosphors can occur from an excited state other than S_1 or T_1. Due to a lack of consistent nomenclature, "dual", "triple", "anomalous emission", and "S_2 emission" have been used to describe emission from a higher energy excited state and are iterchangeable.[3] For the sake of simplicity, "S_2" and "anomalous" will be used to describe emission from a higher energy excited state in this chapter. In this chapter S_2 does mean the second excited state is emissive.

In a recent review, Itoh organized the intramolecular mechanisms responsible for anomalous fluorescence of certain molecules into three types: Schemes A, B, and C (**Fig. 4.1**).[3] Scheme A describes S_2 emission resulting from the thermal population of S_2 from S_1 followed by radiative relaxation to the ground state. In Scheme B, the molecule reaches S_2 by reversible internal conversion (IC) between nearly degenerate vibrational levels of S_1 and S_2 then relaxes radiatively. This mechanism was proposed to exclusively exist in the vapor phase as collisions would prevent IC. Scheme C illustrates S_2 emission resulting from the higher state without IC; the fluorophores are excited into a high energy excited

state not S_1. For the anomalous fluorophores documented in this 2012 review, oscillator strength of S_2 was far greater than that of S_1.[3] Each scheme has a characteristic S_2–S_1 energy gap: this gap is less than 2000 cm^{-1} for Scheme A, less than 3000 cm^{-1} for Scheme B, and 5000-13000 cm^{-1} for Scheme C.

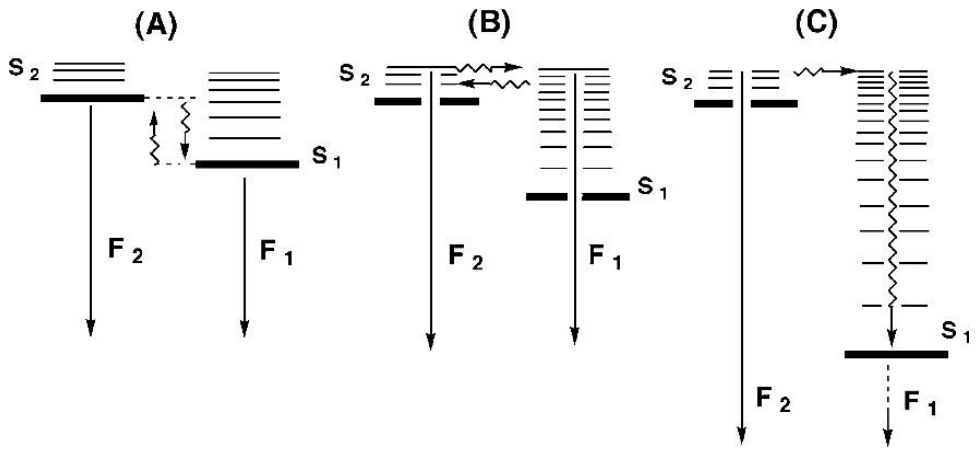


Figure 4.1. These simplified Jablonksi diagrams depict three unique schemes have been proposed to explain the intramolecular mechanisms resulting in anomalous emission. S_1 and S_2 indicate singlet excited states; F_1 stands for fluorescence; and F_2 represents anomalous emission. Taken from reference 3.

Examples of anomalous fluorophores. The anomalous emission of linear polyenes and pyrene can be explained by Scheme A.[3] Scheme B requires collision-free conditions thus the fluorophore must be in the gas phase. Examples of Scheme B fluorophores include vaporized pyrene and large aromatic acenes.[3] Several metalloporphyrins, azulene, and thioketones exhibit fluorescence described by Scheme C.[3] The mechanism responsible for the anomalous emission for some fluorophores may not be able to categorized into one of the schemes described in **Figure 4.1**. The boron

difluorohydrazone (BODIHY) dyes discussed in chapter 2 were also anomalous fluorophores.[4] Anomalous emission has been observed in biological systems such as the enzyme cytochrome P450.[5] Dual emission from S_1 and S_2 emission was observed for the enzyme-bound heme.[5]

Does the Suppression of Kasha's Rule (SOKR) mechanism apply to other compounds exhibiting "anomalous" emission? The anomalous emission enhancement in the aggregate state of the family of BODIHY dyes was explained by the SOKR mechanism.[4] The intramolecular mechanisms described by Schemes A and C share some similarities with the aforementioned SOKR mechanism. Since Scheme B requires collision-free conditions, we can safely rule out SOKR explaining the anomalous emission of Scheme B molecules. The small energy gap between excited states is shared by SOKR and the Scheme A mechanism. For example, the S_1-S_3 energy gap for the unsubstituted BODIHY is 0-5000 cm^{-1} but the energy gap is dependent on the dihedral angle.[4] State crossing occurred when the heterocycle was perpendicular to the phenyl[4]. This range could be characteristic of Scheme A and possibly Scheme C. The population of S_2 by incident excitation energy of Scheme C is also reminiscent of the SOKR mechanism. The direct excitation into S_1 did not lead to the thermal population of $S_{>1}$ for the BODIHY dyes.[4]

Approach for identifying new SOKR molecules. Several Scheme A and C molecules were tested to determine if their anomalous emission could be described by the SOKR mechanism. The anomalous emission, viscosity sensitivity, and barrier to rotation in the emissive state are the mainstays of the SOKR mechanism.[4] The molecules were studied in solvent systems that assessed the dependence of emission intensity on viscosity. A binary solvent mixture, such as the ethylene glycol /glycerol (EG/Gly) mixture described previously, is an attractive solvent system because a wide range of viscosities can be attained.[4, 6] The EG/Gly mixture can produce solutions ranging from 20.8 to 620.7 cP with only small changes to polarity. The combination of polarity and viscosity sensitivity could support a charge transfer transition such as twisted intramolecular charge transfer (TICT) instead of SOKR as the origin for viscosity sensitivity.[6]

As discussed in chapter 2, TICT is another mechanism to describe aggregation induced emission (AIE). Upon photoexcitation, the molecule relaxes from the locally excited (LE) geometry to a twisted complex which is accompanied by an intramolecular charge transfer (TICT).[7] This charge transfer is enhanced in polar solvents as the zwitterionic species is stabilized by the solvent.[7] However, Most molecules will relax from the TICT state through non-radiative pathways.[8] However, radiative relaxation can occur from both the LE state and the TICT state in a dual emission fluorophore as high-energy and low-energy light respestively.[9] In a polar solvent, the LE emission would decrease while (in the case of dual emission) the weaker, red-shifted TICT emission would be enhanced.[9] In the aggregate state, the higher-energy LE state is the emissive state as

molecular motion is restricted; prevention of TICT complex formation leads to emission enhancement.[6]

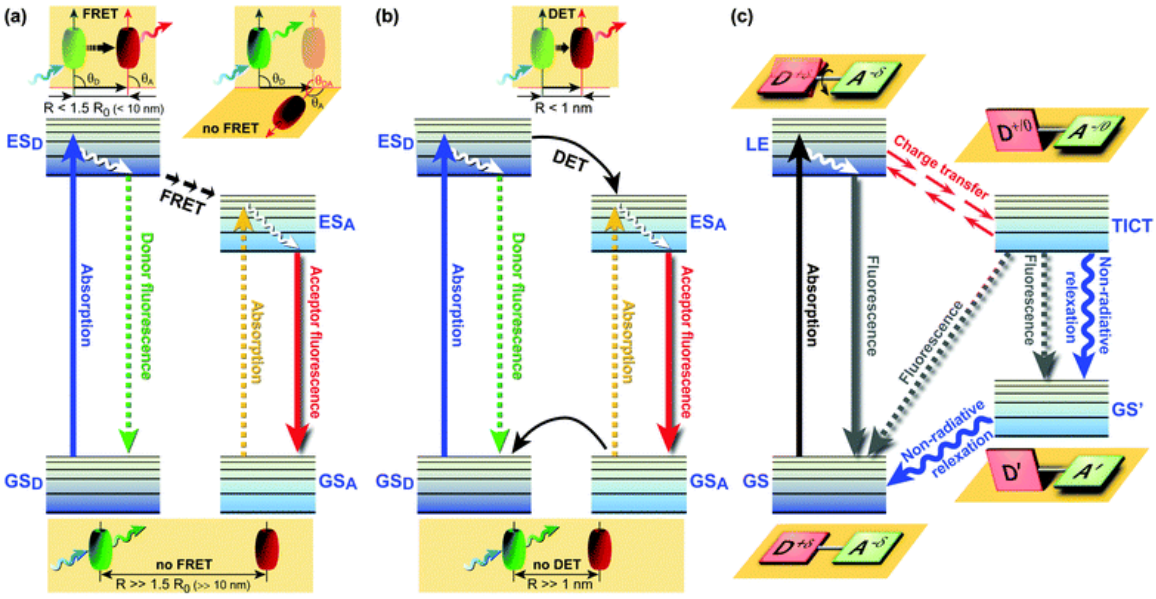


Figure 4.2. Several Jablonksi diagrams describing the photophysical mechanisms of donor-acceptor fluorophores. Förster Resonance Energy Transfer (FRET) and Dexter Energy Transfer (DET) have the donor and acceptor as two separate molecules and are destructive and lead to quenching. FRET requires the donor and acceptor molecules to be within close proximity as well as good overlap between the donor emission and acceptor absorption spectra. DET also requires the donor and acceptor molecules to be within close proximity (or the same molecule). Orbital overlap is required for the two moieties to exchange electrons. The abbreviations are as follows: D: donor; A: acceptor; GS_D: donor ground state; GS_A: acceptor ground state; ES_D: donor excited state; ES_A: acceptor excited state; LE: locally excited; R = effective D–A distance. Figure taken from reference 7.

It is expected that the photophysical properties of molecules from either Scheme A or C could be explained SOKR and would greatly expand the number of fluorescent molecules that fall under this newly proposed and experimentally verified mechanism. If emission is enhanced by increasing viscosity, the fluorophore could be used as a turn-on sensor for viscosity in microenvironments.[10] New SOKR fluorophores could also be used in the fabrication of light emitting diodes (LEDs).

Anomalous fluorophores discussed in this chapter. To better understand the influence of viscosity on anomalous emission, the fluorophore selected from Scheme A was all *trans*-1,6-diphenyl-1,3,5-hexatriene (DPH) and zinc 5,10,15,20-tetraphenyl-21H,23H-porphine (Zn TPP) and azulene (Az) were selected as Scheme C molecules (**Fig. 4.3**).

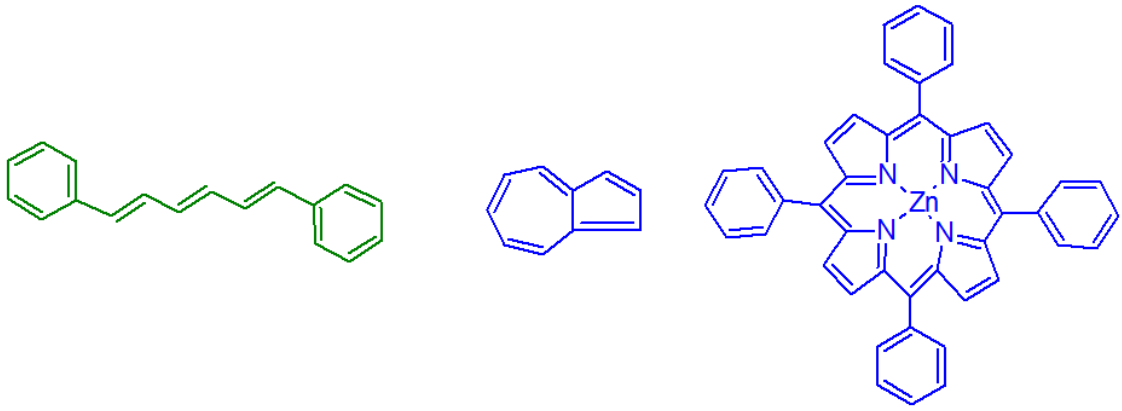


Figure 4.3. The anomalous fluorophores discussed in this chapter: DPH, Az, and Zn TPP (left to right). Scheme A molecule is highlighted in green and the Scheme C molecules are highlighted in blue.

SECTION 4.2. METHODS

Sample Prep. Samples were dissolved in a series of solvents to maximize their miscibility in the EG. DPH was dissolved in DMSO, Zn TPP was dissolved in EtOH, and Az was dissolved in acetone. To prepare the pre-stained EG (PS-EG) solution, an aliquot of 125 μL of the dissolved fluorophore was added to 4.875 mL of EG. PS-EG was mixed with EG and Gly to create the gradient of viscosities as shown in **Table 4.1**. A second set of increasingly viscous solvents were used where a range of 0-50 % Gly was achieved in increments of 10 % steps (**Table C.1**). Samples were sonicated for up to 5 min to ensure homogeneity.

Table 4.1. Solutions to probe viscosity sensitivity of potential SOKR fluorophores.

Sample	**PS-EG** (%)	**EG** (%)	**Gly** (%)	**Viscosity (cP)**
A	10	10	80	620.7
B	10	30	60	283.1
C	10	50	40	124
D	10	60	30	80.8
E	10	70	20	52
F	10	90	0	20.8

UV-vis absorption (Abs) spectroscopy. Abs spectra were collected to optimize parameters for fluorescence spectroscopy. Room-temperature spectra were collected using a Cary 100 Bio spectrophotometer with a scan rate of 600 nm/min and a 1.0 nm data interval. The sample cuvettes were 1.0 cm quartz cells.

Fluorescence spectroscopy. A Photon Technology International QuantaMaster 4 spectrofluorometer outfitted with a LPS-220B lamp power supply, ASOC-10 electronics interface, MD-4000 motor driver control, and a Model 814 photomultiplier detector system was used to collect emission spectra. Data were acquired in 1 nm step sizes, an integration time of 1 s, and slit widths were altered to optimize conditions for the sample. Sample cuvettes were 1.0 cm quartz cells.

SECTION 4.3. RESULTS

It is unclear if all-*trans* DPH emission is influenced by the viscosity of the microenvironment. All-*trans* DPH was dissolved in DMSO then mixed with EG to prepare the PS-EG of all-*trans* DPH. The samples of increasing viscosity were prepared using the solvent cocktails described in **Table 4.1** (viscosity range of 20 to 620 cP). An increase in viscosity led to a decrease in intensity for both S_2 and S_1 emission (**Fig. 4.4**). Spectra were collected in solvent cocktails of increasing viscosity (20-200 cP) to determine if the sample formed aggregates that crashed out of solution in highly viscous environments (viscosity greater than 200 cP) (**Fig. C.1**). This range was selected because the intensity was relatively constant for cocktails with viscosity greater than 200 cP (**Fig. 4.4**). The S_2 and S_1 emission decreased upon increasing viscosity for this viscosity range as well. This phenomenon was observed repeatedly for a range of concentrations (2 to 20 μM). However, all-*trans* DPH inconsistently displayed a positive viscosity sensitivity but this trend was not reproducible and, with $R^2<0.9$, it is unclear if all-*trans* DPH is sensitive to viscosity (**Fig. C.2**). The spectra shown below was the most commonly observed phenomena for all-*trans* DPH: an increase in viscosity led to decrease in both S_1 and S_2 emission.

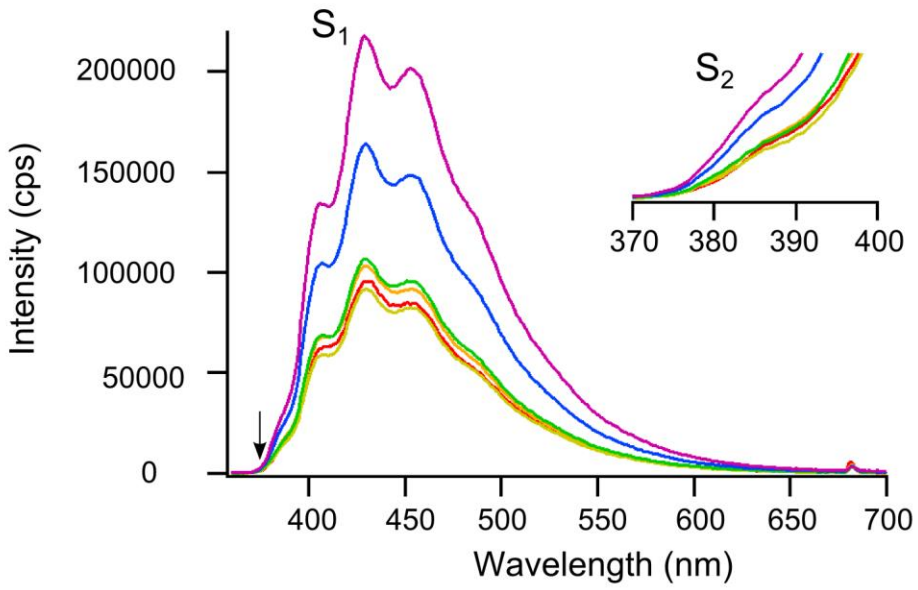


Figure 4.4. Viscosity dependent emission spectra of all-*trans* DPH in mixtures of DMSO, EG, and Gly. The inset is of S_2 emission. Excitation wavelength of 340 nm was used. Traces are the average of 3 scans. Data were acquired with 1 nm slit widths, 1 s integration time, and 1 nm step size. Purple trace is 0 % Gly, blue trace is 20 % Gly, green trace is 30 % Gly, yellow trace is 40 % Gly, orange trace is 60 % Gly, and the red trace is 80 % Gly.

Azulene displayed viscosity sensitivity as S_2 emission decreased in viscous solvents. Az was dissolved in acetone which was used to pre-stain EG; samples were then prepared following **Table 4.1**. As the sample environment became more viscous, the intensity of the S_2 feature decreased (**Fig. 4.5**). Spectra were also collected in less viscous EG/Gly cocktails by Adam Petrucci (20 to 200 cP) (**Fig. C.2**). These overall less viscous solvent cocktails were used determine if aggregates that crashed out of solution formed in solutions with viscosity greater than 200 cP. The intensity of the S_2 emission decreased with increasing viscosity; this fluorophore exhibited negative viscosity dependence.

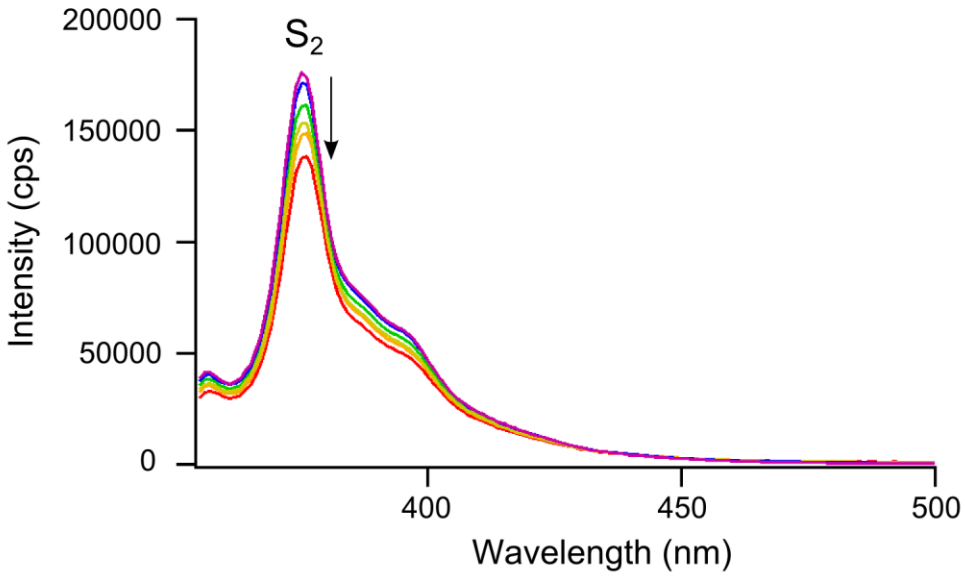


Figure 4.5. Viscosity dependent emission spectra of Az in solutions of acetone, EG, and Gly. An excitation wavelength of 340 nm was used. Traces are the average of 5 scans. Data were acquired with all slit widths set to 1 nm, integration time of 1 s, and step size of 1 nm. Purple trace is 0 % Gly, blue trace is 20 % Gly, green trace is 30 % Gly, yellow trace is 40 % Gly, orange trace is 60 % Gly, and the red trace is 80 % Gly. Raman scattering of solvent cocktail is present at 378 nm.

It is unclear if Zn TPP emission is influenced by the viscosity of the microenvironment. Zn TPP was dissolved in ethanol which was used to pre-stain EG; samples were then prepared following **Table 4.1**. Spectra were collected immediately as the π-stacking of the porphine led to irreversible multimerization in solution.[11] The features of the emission spectrum of Zn TPP arise from two excited states: S_2 emission (428 nm), thermally populated S_1 emission (570 nm), S_1 emission (603 nm), and a lower energy vibrational band of S_1 (656 nm) (**Fig. 4.6**). As viscosity increased, the intensity of both S_1

and S_2 emission decreased. The decreased emission does not arise solely from the viscosity of the microenvironment. The ambiguity is likely due to multiple radiative and non-radiative several relaxation pathways occurring simultaneously.

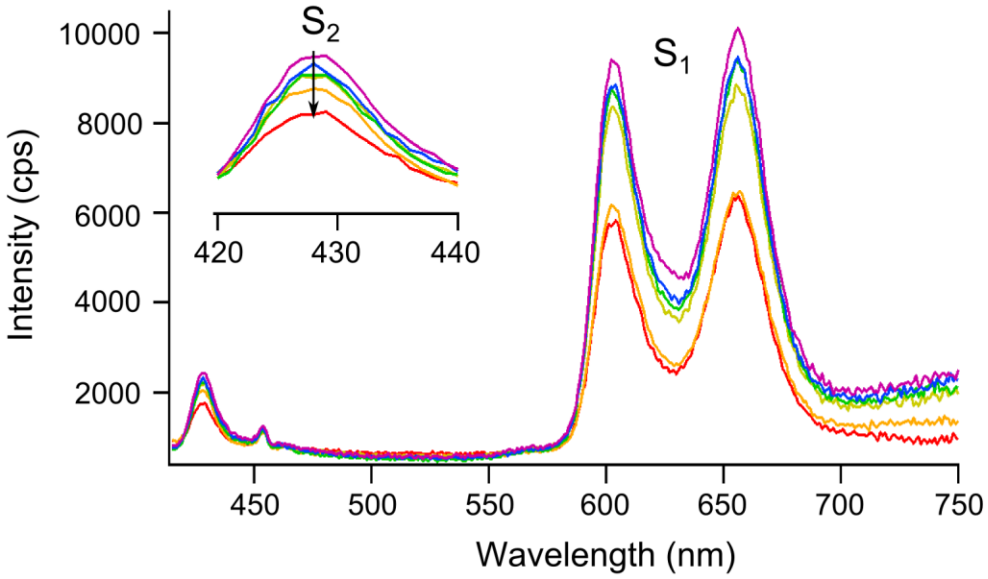


Figure 4.6. Viscosity dependent emission spectra of Zn TPP in solutions of EtOH, EG, and Gly. The inset is of S_2 emission. An excitation wavelength of 400 nm was used to excite the porphine. Traces are the average of 5 scans. Data were acquired with an integration time of 1 sec, step size of 1 nm, slit width of 1 nm. Purple trace is 0 % Gly, blue trace is 20 % Gly, green trace is 30 % Gly, yellow trace is 40 % Gly, orange trace is 60 % Gly, and the red trace is 80 % Gly. Raman scattering of solvent cocktail is present at 453 nm.

SECTION 4.4. DISCUSSION

The influence of microenvironment viscosity on DPH emission is unclear. all-*trans* DPH has two emissive states: S_1 and S_2. Due to a small S_2-S_1 energy gap, the emission of S_2 is a high-energy shoulder on the broad S_1 emission profile. The S_2 emission in all-*trans* DPH is very weak at room temperature and is in thermodynamic equilibrium with the S_1 emission.[12] Diminished S_2 as well as S_1 emission of DPH was observed in the more viscous solvent cocktails (**Fig. 4.7**). However, the linear fit used to derive the negative χ for S_2 emission poorly fit the data. No conclusion can be drawn on the effects of solvent viscosity due to the ambiguous R^2 value. An observed increase in emission intensity as related to increasing viscosity further may be related to the ambiguous R^2 value; this potential AIE behavior was not reproducible.

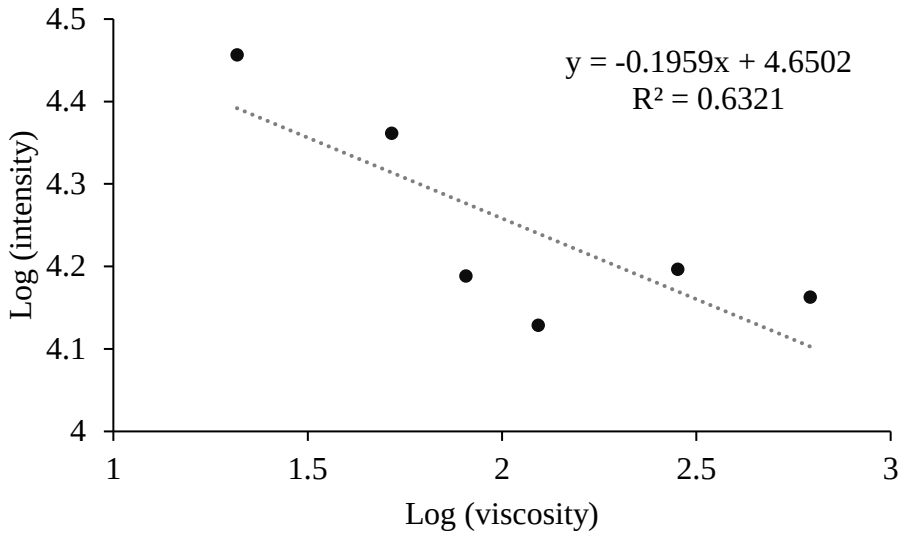


Figure 4.7. Linear fitting of double-logarithmic plot of intensity versus viscosity of DPH in acetone, EG and Gly. Intensity of S_2 emission was recorded at 386 nm. The slope of the best-fit line is equal to the viscosity dependence (χ) and 10^b is equal to the brightness of the fluorophore (*a*).

Because of the R^2 value, a single photophysical event cannot describe the decrease of emission intensity as related to solvent viscosity. Upon exposure to sunlight, this all-*trans* polyene will partially isomerize into mixtures of five accessible *cis*- isomers.[13] Upon exposure to direct sunlight for 1 hr, 35 % of all-*trans* DPH is converted to *cis*- isomers or degraded DPH.[13] Irradiation of all-*trans* DPH at 366 nm leads to the terminal (*ttc*) and central (*tct*) bond isomerization; of the accessible isomers, all-*trans* DPH has the highest quantum yield.[14]

The excitation wavelength used to excite all-*trans* DPH for fluorescence spectroscopy was 340 nm; the excitation wavelength may be causing isomerization of DPH

during the time course of the experiment. It was initially proposed that the dramatic decrease in intensity of the emission spectra was due to photobleaching. The instrument was modified to minimize photobleaching (smaller excitation slit widths and wider emission slit widths). With these modifications to the instrument, emission intensity still decreased during subsequent scans of the sample. This decreased intensity could be due to photoisomerization to isomers with lower quantum yields.[14]

Complete conversion to single DPH isomer has not been reported; therefore, a several isomers are present in solution following photoisomerization. These isomers can be separated by HPLC and monitored by ^{1}H NMR.[14-15] The *cis*- isomers of DPH less stable than all-*trans* DPH and may degrade at room-temperature.[13] The emission spectra of DPH are not excitation wavelength independent thus the component of a partial *cis*- isomer (s-*cis* DPH) has be elucidated from the difference spectra.[16] Two emission spectra are collected of DPH and the excitation wavelength is either 355 nm; the difference between these two spectra is a low intensity and red-shifted emission band attributed to s-*cis* DPH.[16]

Azulene is an unlikely SOKR molecule. Az was selected due to its deep history as the first published anomalous fluorophore.[2] This major emissive state of this fluorophore is S_2 while S_1 is a weakly emissive state; the quantum yield of S_2 emission is about 4-orders of magnitude greater than S_1 emission when in solution.[17] S_1 emission occurs in the near-IR (740-940 nm).[18] However, the noise is too great beyond 650 nm due to the limitations of the R928 photomultiplier tube (Hamamatsu Photonics) in the detector.

Of the three dyes investigated in this chapter, viscosity had the least influence on the anomalous emission of Az. It is likely that an ACQ phenomenon is responsible for the

negative χ of S_2 emission in Az (**Fig. 4.8**). The goodness of fit was sufficient to unambiguously assign the negative viscosity sensitivity of azulene (R^2>0.95). ACQ is very common for planar aromatic fluorophores.[19] The formation of excimers (dimerization of two excited fluorophores) or exciplexes (dimerized of an excited and ground state fluorophore) in planar organic fluorophores leads to non-radiative relaxation pathways thus decreasing emission.[20] The emission enhancement described by the SOKR mechanism is contingent on aggregation induced emission Because of the negative viscosity sensitivity of S_2 emission, it is unlikely that the anomalous emission of azulene can be explained by the SOKR mechanism. While not a SOKR fluorophore, azulene could be used as a turn-off sensor for viscous environments.

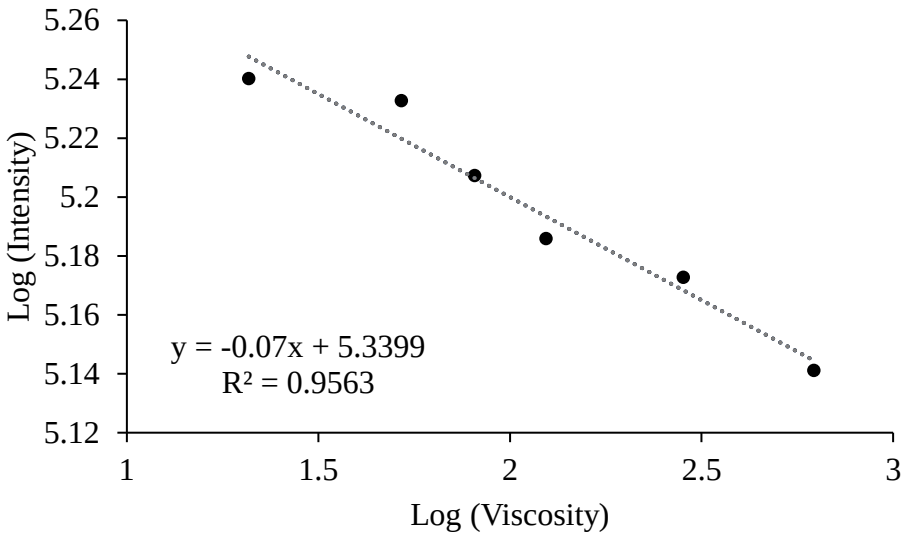


Figure 4.8. Linear fitting of double-logarithmic plot of intensity versus viscosity of azulene in acetone, EG and Gly. Intensity of S_2 emission was recorded at 376 nm. The slope of the best-fit line is equal to the viscosity dependence (χ) and 10^b is equal to the brightness of the fluorophore (*a*).

Zn TPP as a potential SOKR molecule. Zn TPP is emissive from three states: S_2, S_1, and T_1. The porphines can form π-stacking aggregates that result in quenching of emission (an inherent downfall of using these molecules as fluorescent probes).[11] This aggregation was monitored spectroscopically. The intensity of the Soret peak (420 nm) in the Abs spectrum and the emission intensity was reduced (**Fig. C.3**). The π-stacking was accompanied by a color change from purple to green. To circumvent this, samples were scanned within 30 minutes of preparation. The Stokes shift of Zn TPP is very small (115 ± 15 cm^{-1}) thus 1 nm slit widths were used to prevent reabsorption and μM concentrations were used to minimize the stacking interaction.[21]

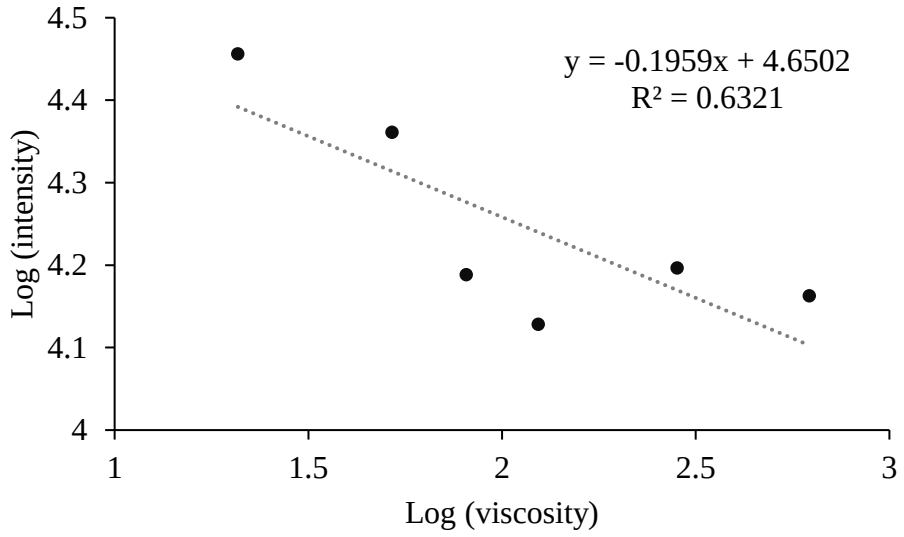


Figure 4.9. Linear fitting of double-logarithmic plot of intensity versus viscosity of Zn TPP in EtOH, EG and Gly. Intensity of S_2 emission was recorded at 328 nm. The slope of the best-fit line is equal to the viscosity dependence (χ) and 10^b is equal to the brightness of the fluorophore (*a*).

SECTION 4.5. CONCLUSIONS AND FUTURE WORK

Azulene did not exhibit viscosity dependent emission. Viscosity dependence was investigated to determine if the dyes were AIE luminogens. The negative viscosity dependence of anomalous emission from Az indicated that the SOKR mechanism could not describe the photophysical properties of this fluorophore. If these luminogens were SOKR fluorophores, a positive χ for S_2 emission would be expected.

Ambiguous solvent sensitivity of DPH requires further investigation. On two separate occasions, a positive viscosity sensitivity was observed for DPH (concentration of 2 μM and 20 μM). The preparation of the sample appears to have been consistent but this phenomenon was not reproducible. Revision of experimental protocol is required due to the lack of reproducibility of the AIE phenomenon coupled with the low R^2 value for negative viscosity sensitivity.

The ambiguity of the viscosity sensitivity likely arose from many competing photochemical processes. The decrease in emission could be from static quenching (π-stacking in the ground state).[19] Luminogens with large aromatic components as well as rod-shaped luminogens are often plagued by ACQ. Several DPH isomers are present in solution due to the formation of *cis*- isomers following excitation.[13] To understand the viscosity dependence on all-*trans* DPH, the contribution of *cis*- isomers will be determined spectroscopically.

The difference spectra can provide information pertaining to the contribution of other isomers emission spectra. The Abs spectra of the isomers are markedly similar but the spectrum of s-*cis* DPH has a low energy shoulder and less vibrational structure.[22] The photoisomerization cannot be avoided as the excitation wavelength will lead to *cis*-isomers. The contribution of *cis*- isomers can be determined from the difference emission spectra, HPLC, or ^{1}H NMR.[13-14, 23] Due to the high-energy excitation wavelength, it is unlikely that photoisomerization can be prevented thus the contribution of *cis*- isomers of DPH must be investigated.

Ambiguous solvent sensitivity of Zn TPP requires further investigation. Static quenching from π-stacking in the ground state led to spectral interference. Due to the spectral overlap of the Soret band in the Abs spectrum and S_2 emission and proximity of monomers, FRET is a likely occurring during π-stacking.[7] It is likely that the constructive quenching phenomena from π-stacking and FRET led to the ambiguous R^2 value (**Fig. 10.9**). Binary mixtures of short chained alcohols can be used to monitor the viscosity sensitivity while preventing π-stacking. The longer chained alcohols are more viscous (0.54 cP, methanol; 1.07 cP, ethanol; 1.96 cP, propanol at 25 °C) and less polar (relative polarity: 0.72, methanol; 0.65, ethanol; 0.617, propanol).

SOKR may exclusively describe the emission of AIE molecular rotors. It may be that SOKR is specific to molecular rotors. The barrier to rotation in the excited state of the BODIHY dyes was correlated to their viscosity sensitivity.[4] Thus, it is worth investigating the influence of viscosity on anomalous emission for some molecular rotors in the future. If the emission of a future molecule is determined to be viscosity dependent and polarity independent, the photophysical properties would be modeled by density functional theory (DFT) and time-dependent DFT (TDDFT). The SOKR mechanism described the anomalous emission of AIE luminogens arising from restriction of rotation into the S_1 state.[4] The DFT and TDDFT modelling of the molecules would be used to predict the $\Delta E(S_{>1} - S_1)$ energy gap as well as the barrier to rotation in the excited state.

SECTION 4.6.REFERENCES

1. Kasha, M., *Disc. Faraday Soc.* **1950,** *9*, 14-19.

2. Beer, M.; Longuet-Higgins, H. C., *J. Chem. Phys.* **1955,** *23*, 1390.

3. Itoh, T., *Chem. Rev.* **2012,** *112*, 4541-4568.

4. Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., *Nat. Chem.* **2017,** *9*, 83-87.

5. Barnaba, C.; Humphreys, S. C.; Barden, A. O.; Jones, J. P.; Brozik, J. A., *J. Phys. Chem. B* **2016,** *120*, 3038-3047.

6. Haidekker, M. A.; Brady, T. P.; Lichlyter, D.; Theodorakis, E. A., *Bioorg. Chem.* **2005,** *33*, 415-425.

7. Sasaki, S.; Drummen, G. P. C.; Konishi, G., *J. Mater. Chem. C* **2016,** *4*, 2731-2743.

8. Abdel-Mottaleb, M. S. A.; Loutfy, R. O.; Lapouyade, R., *Photochem. Photobiol.* **1989,** *48*, 97-93.

9. Haidekker, M. A.; Theodorakis, E. A., *J. Mater. Chem. C* **2016,** *4*, 2707-2718.

10. Haidekker, M. A.; Theodorakis, E. A., *J. Biol. Eng.* **2010,** *4*, 11.

11. Guo, B.; Cai, X. L.; Xu, S. D.; Fateminia, S. M. A.; Liu, J.; Liang, J.; Feng, G. X.; Wu, W. B.; Liu, B., *J. Mater. Chem. B* **2016,** *4*, 4690-4695.

12. Itoh, T.; Kohler, B. E., *J. Phys. Chem.* **1986,** *91*, 1760-1764.

13. Lunde, K.; Zechmeister, L., *J. Am. Chem. Soc.* **1953,** *76*, 2308-2313.

14. Saltiel , J.; Krishnamoorthy , G.; Huang , Z.; Ko , D.-H.; Wang, S., *J. Phys. Chem. A* **2003,** *107*, 3178-3186.

15. Mesganaw, T.; Im, G. Y.; Garg, N. K., *J. Org. Chem.* **2013,** *78*, 3391-3393.

16. Saltiel, J.; Sears Jr., D. F.; Sun, Y.-P.; Choi, J.-O., *J. Am. Chem. Soc.* **1992,** *114*.

17. Wagner, B. D.; Tittelbach-Helmrich, D.; Steer, R. P., *J. Phys. Chem.* **1992,** *96*, 7904-7908.

18. Orenstein, M.; Kimel, S.; Speiser, S., *Chem. Phys. Lett.* **1978,** *58*, 582-585.

19. Mei, J.; Leung, N. L.; Kwok, R. T.; Lam, J. W.; Tang, B. Z., *Chem. Rev.* **2015,** *115*, 11718-11940.

20. Hong, Y.; Lam, J. W.; Tang, B. Z., *Chem. Soc. Rev.* **2011,** *40*, 5361-5388.

21. Karolczak, J.; Kowalska, D.; Lukaszewicz, A.; Maciejewski, A.; Steer, R. P., *J. Phys. Chem. A* **2004,** *108*, 4570-4575.

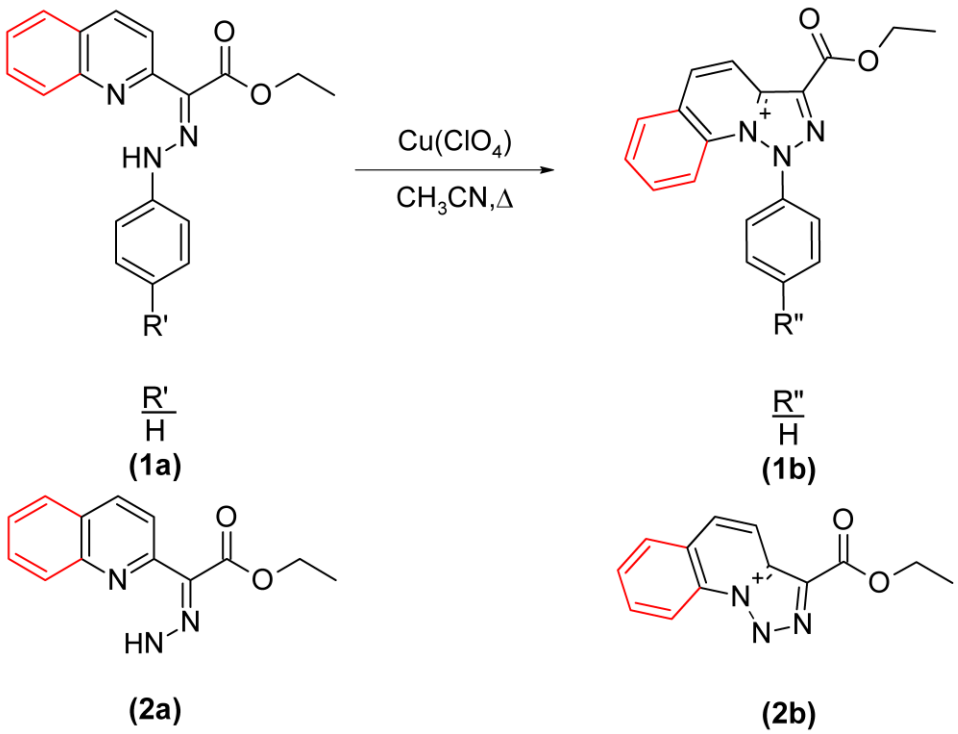
22. Turek, A. M.; Saltiel, J.; Krishna, T. R.; Krishnamoorthy, G., *J. Phys. Chem. A* **2012,** *116*, 5353-5367.

23. Zechmeister, L., *Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes*. Academic Press Inc Publishers: New York, 1962; p 266.

CHAPTER 5: MODELING THE LARGE STOKES SHIFTS OF TRIAZOLOPYRIDINIUM AND TRIAZOLOQUINOLINIUM DYES

SECTION 5.1 INTRODUCTION

[1,2,3]triazolo[1,5-a]pyridinium dyes are water-soluble fluorophores. The Aprahamian group (Dartmouth College, Hanover, NH) published an oxidative cyclization synthetic route for the triazolopyridinium (TOP) dyes in 2012.[1] The TOP dyes were synthesized as a monomer (R' = R" = H) and dimer (R' = open triazole moiety, R" = cyclized triazole moiety) (**Scheme 5.1**). This copper-mediated cyclization was expanded to include the [1,2,3]triazolo[1,5-a]quinolinium (TOQ) dyes shortly thereafter.[2] These TOQ dyes were also synthesized as a monomer and dimer (**Scheme 5.1**).



Scheme 5.1. Synthesis of [1,2,3]triazolo[1,5-*a*]pyridinium (1b) and [1,2,3]triazolo[1,5-a]quinolinium (2b) from the preceding hydrazone. The aerobic Cu(II) mediated cyclization occurs in acetonitrile at 60 °C (65-75% yield). Scheme adapted from references 1 and 2.

Potential applications for TOP and TOQ dyes. The market for water-soluble fluorophores is vast and includes many biological applications where aqueous conditions are required for the analyte. Typically, large conjugated π-systems of fluorophores lead to poor water solubility. The novelty of these TOP dyes partially lies in their solubility properties as well as their blue emission. However, the expansion of the π-system from pyridinium to quinolinium led to poor water solubility.[2] These dyes could be used in the light-emitting layer in a white or blue organic light emitting diode (OLED). There are two approaches to creating white OLEDs: (1) single luminogen with broad emission or (2) combination of red, green, and blue luminescent components.[3]

The broad emission of the TOQ and TOP dyes could be used to fabricate white OLEDs. Structural modifications could be made to the triazole dyes to widen the emission band to include lower energy red emission. The broader emission profile of the TOP monomer compared to the dimer is due to distortions in the excited state arising from a large post-excitation charge transfer.[1] These dyes could also be used as stand-alone blue OLEDs; the four triazole dyes emit green-blue to deep-blue light as defined by the Commission Internationale de L'éclairage (CIE) 1931 color space.[2] Efficient blue emitters have been elusive due to stability and efficiency problems.[4] They could be used in combination with dyes with green and red emission to produce white light.

One metric for fluorescent molecules is quantum yield (Φ_F) which is the ratio of photons absorbed to photons emitted as fluorescence. A larger Φ_F is indicative of a brighter fluorophore. The Φ_F of the TOP dyes in water increased from 0.05 ± 0.02 for the monomer to 0.30 ± 0.02 for the dimer.[1] This significant increase in Φ_F as well as the red-shifted for the TOP dimer implies this dimer act as a single chromophore instead of two chromophores connected by a phenyl linker.[2] The Φ_F was 0.04 and 0.11 for the TOQ monomer and dimer in methanol and acetonitrile respectively.[2] Due to the increased Φ_F accompanied by a red-shift in the UV-vis absorption (Abs) spectra, the TOQ dimer is expected to act as a single chromophore instead of two subunits.

Another attribute that plagues inefficient fluorophores is a small Stokes shift. The Stokes shift is the difference in energy of the band maxima in the Abs spectrum and the emission spectrum. Due to a relaxation events following excitation, the $\Delta E_{0\rightarrow1}$ is greater than $\Delta E_{1\rightarrow0}$. The immediate solvent molecules as well as the fluorophore relax after excitation; significant changes to fluorophore's geometry as well as the surrounding solvent molecules will result in a large Stokes shift. A small Stokes shift leads to high instances of self-quenching due to reabsorption. In some cases, and at the cost of overall brightness, a dilute solution of the fluorophore (nM) can be used avoid self-quenching of fluorophores with small Stokes shifts.[5]

The TOP and TOQ dyes discussed in this chapter have large Stokes shifts. Large Stokes shifts are rare for organic luminogens and more commonly attributed to inorganic complexes.[5] The large Stokes shift and water solubility make these dyes very attractive for many applications such as biological imaging. A computational method to predict the Stokes shift of the triazole-based fluorophores will be discussed in this chapter. This methodology provides more information pertaining to the excited state of the fluorophore. A systematic approach to predict the Stokes shift can be used to design more efficient fluorophores.

SECTION 5.2 METHODS

Density functional theory (DFT) modeling of the Abs spectra of TOP and TOQ dyes. The photophysical properties of the TOP monomer and dimer were modeled using the PBE functional in 2012.[1] To identify the best functional for modeling the TOP and TOQ dyes, four functionals were investigated for their ability to accurately predict the Abs spectrum of the TOP monomer and dimer and the TOQ monomer and dimer. The functionals were from two classes: GGA (PBE and BLYP) and hybrid (PBE0 and B3LYP).[6] The other triazole-based dyes were built in ArgusLab from the crustal structure of the TOP monomer (CCDC number 887676).

All calculations discussed in this chapter were completed using ORCA 3.0.0. software on the 380 node IBM Bluemoon cluster at the Vermont Advanced Computing Core.[7] Density functional theory (DFT) calculations used a TZVP basis set and tight SCF convergence criteria. The optimizations of TOP and TOQ were performed in the presence of the COSMO continuum solvation model for water ($\varepsilon = 80.4$, $\eta = 1.33$). Time-dependent DFT (TDDFT) was used to predict the energies and intensities of the Abs spectra; 20 excited states within the expansion space of 120 vectors were used for these water models of each functional. The Abs spectra were created using the orca_mapspc utility program. A full-width at half-maximum of 2500 cm^{-1} was selected to mimic the experimental absorption spectra of these dyes.

Predicting Stokes shift of TOP and TOQ dyes. The optimized geometry for the ground state was used as a starting point for the TDDFT determined S_1 geometry. When modeling the larger molecules with hybrid functionals, RIJCOSX was used to alleviate the computational costs.[8] The excited state geometry was then solvated using the COSMO continuum solvation model in the subsequent TDDFT calculation to determine the $S_1 \rightarrow S_0$ energy gap at the S_1 geometry. The TOP dyes were solvated with water ($\varepsilon = 80.4$, $\eta = 1.33$), TOQ monomer was solvated with methanol ($\varepsilon = 32.63$ and $\eta = 1.329$) and the TOQ dimer was solvated with acetonitrile ($\varepsilon = 36.6$ and $\eta = 1.344$). The Stokes shift was calculated as the energy difference between the first excited state at the ground state and excited state geometries.

SECTION 5.3 RESULTS

Verification of functionals. The TOP and TOQ dyes were modeled by the GGA and hybrid-GGA functionals (**Figures D.1-D.4**). The predicted intensity of the lowest energy feature in the Abs spectra of the TOP and TOQ dyes was used to rank the functionals ability to model the excited state electronic structure of the dyes at the ground state geometry (**Table 5.1** and **5.2**). The predicted energy of the Abs feature provides information about the HOMO-LUMO gap while the intensity provides information pertaining to the shape of the donor and acceptor orbitals. An accurately predicted intensity ensures that the transition being modeled is the measured transition. For the four dyes, the most accurate functional was PBE. However, the B3LYP functional modeled the TOQ dyes more accurately while the most accurate predicted Abs spectra of the TOP dyes was with BLYP. The hybrid functionals overestimated the intensity of the TOP dimer by over 30,000 M^{-1} cm^{-1}.

Table 5.1. TDDFT-predicted Abs spectra of TOP dyes compared to the experimental Abs spectra. TOP dyes were solvated with water using the COSMO continuum solvation model.

TOP monomer				
	E (cm^{-1})	**ε (M^{-1} cm^{-1})**	**Error E**	**Error ε**
Experimental	33100	11100	-	-
PBE	26600	8300	-6500	-2800
BLYP	26000	7100	-7100	-4000
PBE0	35900	23200	2800	12100
B3LYP	33630	16310	530	5210
TOP dimer				
	E (cm^{-1})	**ε (M^{-1} cm^{-1})**	**Error E**	**Error ε**
Experimental	32200	22100	-	-
PBE	26400	31400	-5800	9300
BLYP	26200	26900	-6000	4800
PBE0	32800	59700	600	37600
B3LYP	32800	56400	600	34300

Table 5.2. TDDFT-predicted Abs spectra of TOQ dyes compared to the experimental Abs spectra. Dyes were solvated with water using the COSMO continuum solvation model.

TOQ monomer				
	E (cm^{-1})	**ε (M^{-1} cm^{-1})**	**Error E**	**Error ε**
Experimental	29300	8100	-	-
PBE	28000	4900	-1300	-3200
BLYP	27100	3400	-2200	-4700
PBE0	33200	12700	3900	4600
B3LYP	31500	9200	2200	1100
TOQ dimer				
	E (cm^{-1})	**ε (M^{-1} cm^{-1})**	**Error E**	**Error ε**
Experimental	28800	16400	-	-
PBE	22100	9800	-6700	-6600
BLYP	20600	6600	-8200	-9800
PBE0	31100	29400	2300	13000
B3LYP	29700	24600	900	8200

The Stokes shifts of the TOP and TOQ dyes. The Stokes shift was calculated as the TDDFT predicted energy difference between the ground and first excited states at the S_0 and S_1 geometries. The GGA functionals underestimated the Stokes shifts of the monomer while significantly overestimating the shift of the dimer (**Table 5.3**). Conversely, the hybrid-GGA functional PBE0 underestimated the Stokes shifts of the dimer while overestimating the shift of the monomer. B3LYP predicted that same Stokes shift for the TOP monomer and dimer. The Stokes shifts of the TOQ dyes were overestimated by every functional (**Table 5.4**). However, PBE0 was able to model the relative ranking of Stokes shifts: with TOP monomer having the largest Stokes shift, then TOQ monomer, TOP dimer, and TOQ dimer. PBE0 most accurately predicted the Stokes shifts of the four triazole dyes regarding the absolute average error and predicting the trend of the shifts.

Table 5.3. Predicted Stokes shift of TOP dyes. Experimentally determined Stokes shifts from reference 2. Dyes were solvated with water using the COSMO continuum solvation model.

TOP monomer		
	Stokes shift (cm^{-1})	**Error**
Experimental	11200	-
PBE	6400	4800
BLYP	8100	3100
PBE0	17900	-6700
B3LYP	4800	6400
TOP dimer		
	Stokes shift (cm^{-1})	**Error**
Experimental	7800	-
PBE	19500	-11700
BLYP	20200	-12400
PBE0	4100	3700
B3LYP	4800	3000

Table 5.4. Predicted Stokes shift of TOQ dyes. Experimentally determined Stokes shifts from reference 2. Dyes were solvated with methanol and acetonitrile using the COSMO continuum solvation model.

TOQ monomer		
	Stokes shift (cm^{-1})	**Error**
Experimental	9200	-
PBE	12500	-3300
BLYP	11300	-2100
PBE0	16000	-6800
B3LYP	16000	-6800
TOQ dimer		
	Stokes shift (cm^{-1})	**Error**
Experimental	7000	-
PBE	17100	-10100
BLYP	17400	-10400
PBE0	9900	-2900
B3LYP	10800	-3800

Excited state geometries. As PBE0 best modeled the TOP and TOQ Stokes shifts, the geometries were investigated to determine the changes in the excited state. The gas-phase optimized calculations predicted that S_1 geometry was markedly different from the ground state geometry for both the TOP and TOQ monomer (**Fig. 5.2**). Conversely, the TOP and TOQ dimers only experienced minimal changes in their configurations following excitation (**Fig. D.5**). This significant geometry change for the monomers explains the larger Stokes shifts of these two dyes over the dimers.

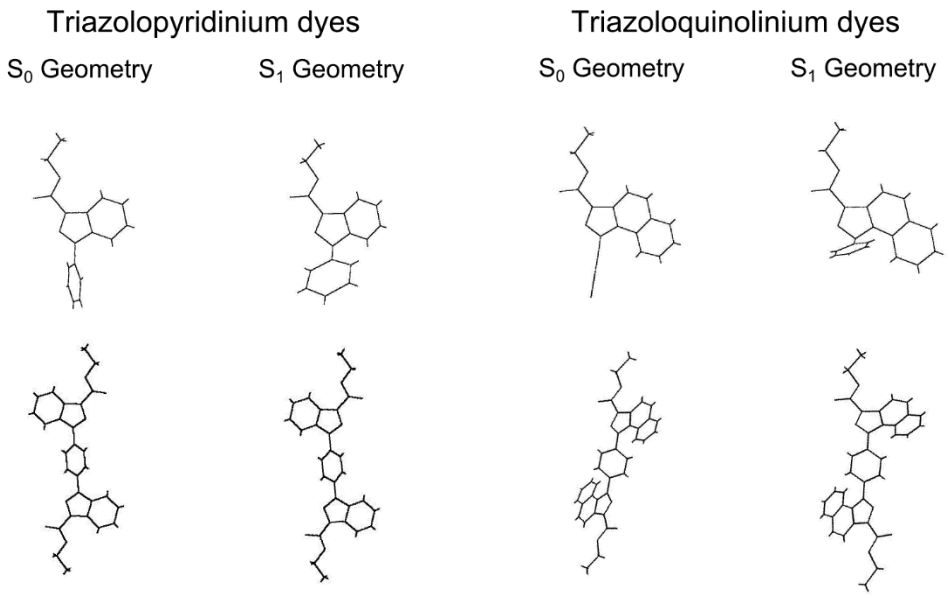


Figure 5.1. The S_0 (first and third column) and S_1 (second and fourth column) geometries of the monomers (top row) and dimers (bottom row) of the TOP dyes (left columns) and TOQ dyes (right columns).

SECTION 5.4 DISCUSSION

PBE most accurately predicted Abs spectra. The most accurate modeling of the donor and acceptor orbitals was by the GGA functional PBE. PBE, B3LYP, and PBE0 modeled the trend intensity of the lower energy feature of the four dyes correctly where TOP dimer has the most intense feature followed by TOQ dimer, TOP monomer, and TOQ monomer. PBE was not only shown to accurately model the intensity trend, but also had the smallest average absolute error (**Table 5.5**).

Table 5.5. The absolute error of the functionals predicting the Abs intensity of the TOP and TOQ dyes.

Functional	**Average absolute error (M^{-1} cm^{-1})**
PBE	5500
BLYP	5800
PBE0	16800
B3LYP	12200

PBE0 most accurately predicted the Stokes shift. The relaxation of the solvent and fluorophore was severely overestimated for both TOQ dyes by all functionals. Only PBE0 correctly predicted that the dimer would have a smaller Stokes shift compared to the monomer. The most accurate function for predicting the Stokes shift regarding the average absolute error of the energy gap was PBE0 (**Table 5.6**). It is expected that the most accurate modeling of the relaxation following excitation was by the hybrid-GGA functional PBE0.

Table 5.6. The absolute error of the functionals predicting the Stokes shift of the TOP and TOQ dyes.

Functional	**Average absolute error (cm^{-1})**
PBE	7500
BLYP	7000
PBE0	5000
B3LYP	5400

The Stokes shift is due to relaxation of the fluorophore as well as the surrounding solvent. Comparing the ground state and excited state geometries as predicted by PBE0 provides information on the relaxation of the fluorophore without solvent. The differences between the ground and excited state geometries were more pronounced in the monomers (**Figs. 5.1** and **D.5**).

Changes to the geometry following excitation were quantified by monitoring changes to the dihedral angle as well as changes to bond lengths in the heterocycle (**Table 5.7**). The bond lengths were measured in ArgusLab (Planaria Software). The three bonds lengths selected were N—N (bond A), N—C (bond B), and N—N (bond C) of the triazole heterocycle (**Fig. 5.2**). The greatest Stokes shift was predicted for the TOP monomer and this dye exhibited unique changes to the excited state geometry. Elongation of the N—C bond linking the phenyl to the heterocycle was elongated following excitation only for the TOP monomer. For all dyes, except the TOP monomer, the two N—N bonds were elongated after the relaxation in the excited state.

Table 5.7. Structural changes to geometry following relaxation that lead to large Stokes shifts for the TOP and TOQ dyes.

	State	Bond lengths with a single triazole moiety (Å)			Bond lengths of the second triazole moiety (Å)			Dihedral angle (°)
		A	B	C	A	B	C	
TOP monomer	S_0	1.45	1.40	1.42				68
	S_1	1.30	1.43	1.36				-14
TOQ monomer	S_0	1.30	1.44	1.37				98
	S_1	1.44	1.40	1.42				14
TOP dimer	S_0	1.31	1.42	1.36	1.31	1.42	1.36	54, 54
	S_1	1.34	1.39	1.39	1.31	1.42	1.36	39, -56
TOQ dimer	S_0	1.31	1.43	1.37	1.31	1.43	1.37	54, 38
	S_1	1.40	1.38	1.41	1.32	1.40	1.37	-16, 31

The torsion angle was measured to monitor the rotation around of the phenyl around the heterocycle (**Fig. 5.2).** For both monomers, the phenyl rotated over 80 ° from its starting position in the ground state geometry after excitation. The rotation of one of the heterocycles was significantly greater than the other heterocycle for the TOP dimer as well as the TOQ dimer. For the TOP dimer, one heterocycle rotated 110° from its starting position while the other heterocycle rotated 15°. For the TOQ dimer, one heterocycle rotated 70° from its starting position while the other heterocycle rotated 7°; the TOQ dimer was predicted to have the smallest Stokes shift and this is exemplified by the minimal change to the geometry following excitation.

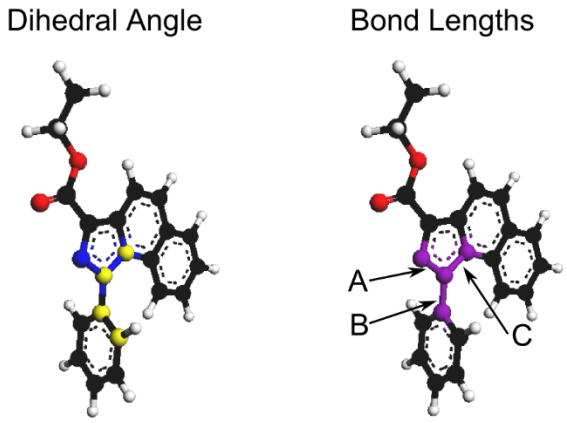


Figure 5.2. Investigation into changes of the geometries in the ground and excited states were determined in ArgusLab (Planaria software). (**Left**) The dihedral angle was measured between the plane of the heterocycle and the phenyl using the four highlighted atoms: N—N—C—C. (**Right**) The three bond lengths (A, B, and C) of the triazole used to compare the ground state and excited state are highlighted in purple. Carbon atoms are black, nitrogen atoms are blue, oxygen atoms are red, and hydrogen atoms are white.

SECTION 5.5 CONCLUSIONS AND FUTURE WORK

Firstly, the predicted Abs spectra were used to rank the accuracy of the selected functionals. An accurately predicted intensity would determine that the functional was predicting an accurate shape for the donor and acceptor orbitals. The most accurate functional to predict the intensity of the lowest energy Abs feature was PBE. All the functionals discussed in this chapter, PBE, BLYP, PBE0, and B3LYP, overestimated the Stokes shift of the TOQ dyes and erroneously predicted the TOP dyes' Stokes shift. Experimentally, the Stokes shift of the monomer was greater than that of the corresponding dimer.[2] The only functional to predict this trend was PBE0. The significant structural changes between the predicted ground state and excited state geometries were correlated to large Stokes shifts.

To test whether that PBE0 provides to most accurate model of the TOP and TOQ dyes, other functional classes will be tested. All classes of DFT functionals calculate the electronic structure of the ground state by relying on different approximations of the exchange and correlation energy (E_{XC}). The GGA functionals (PBE and BLYP) calculate the electron density as a point as well as a gradient.[9] The meta-GGAs (such as TPSS) approximate E_{XC} with Kohn-Sham kinetic energy densities or density Laplacian.[10] Hybrid functionals are not purely based on DFT as they incorporate a Hartree-Fock (HF) integral at a fixed fraction of the exchange energy.[9] Range-corrected GGA functionals separate the electron repulsion into long-range short-range interactions which are calculated with HF and GGA exchange energy respectively.[11] Like range-corrected GGA functionals, range-

corrected hybrid functionals are hybrids parametrized with HF exchange.[12] Lastly, a double hybrid adds a perturbation term to HF exchange energy and GGA E_{XC}.[13]

New functionals with different treatments of the exchange and correlation energies may model the triazole dyes with greater accuracy. The meta-GGA functional (TPSS) was used to model the TOQ dyes and will now be used to model the TOP dyes.[10a] The TPSS-predicted Abs spectra was inaccurate so it was not used to model the TOP dyes initially (**Table 5.8**). However, the predicted intensity of the lowest energy feature by TPSS was more accurate than what was predicted by PBE0 ergo it is worth investigating how well this meta-GGA functional can model the TOP dyes.

Table 5.8. The lowest energy feature of the TPSS-predicted Abs spectra of the TOQ dyes.

	E (cm^{-1})	**ε (M^{-1} cm^{-1})**	**Error E (cm^{-1})**	**Error ε (M^{-1} cm^{-1})**
TOQ monomer	27400	4300	-1900	-3800
TOQ dimer	21400	5700	-7400	-10700

A wider spectrum of functionals will be used to predict the Stokes shifts of the TOP and TOQ dyes. The selected functionals are a hybrid meta-GGA (TPSSh), range corrected GGA (ωB97), range corrected hybrid (ωB97X), and a double hybrid (B2-PLYP) functional.[10b, 13-14]. To date, the TOP monomer and dimer have been investigated PBE DFT and TDDFT and TOQ monomer was investigated with the semi-empirical MOPAC/MNDO calculations.[1-2]

SECTION 5.6 REFERENCES

1. Su, X.; Liptak, M. D.; Aprahamian, I., *Chem. Commun. (Camb)* **2013,** *49*, 4160-4162.

2. Carboni, V.; Su, X.; Qian, H.; Aprahamian, I.; Credi, A., *Chem. Photo. Chem.* **2017,** *1*, 222-229.

3. Wang, Q.; Ma, D., *Chem. Soc. Rev.* **2010,** *39*, 2387-2398.

4. Hong, M. K.; Ravva, M. K.; Winget, P.; Bredas, J. L., *Chem. Mater.* **2016,** *28*, 5791-5798.

5. Martin, A.; Long, C.; Forster, R. J.; Keyes, T. E., *Chem. Commun. (Camb)* **2012,** *48*, 5617-5619.

6. (a) Perdew, J. P.; Burke, K.; Ernzerhof, M., *Phys. Rev. Lett.* **1996,** *77*, 3865-3868; (b) Lee, C.; Yang, W.; Parr, R. G., *Phys. Rev. B* **1987,** *37*, 785-789; (c) Kim, K.; Jordan, K. D., *J Chem Phys* **1994,** *98*, 10089-10094; (d) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., *J. Chem. Phys.* **1994,** *98*, 11623-11627; (e) Perdew, J. P.; Emzerhof, M.; Burke, K., *J. Chem. Phys.* **1996,** *105*, 9982-9985.

7. Neese, F., *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2*, 73-78.

8. Petrenko, T.; Kossmann, S.; Neese, F., *J. Chem. Phys.* **2011,** *134*, 054116.

9. Burke, K.; Wagner, L. O., *International Journal of Quantum Chemistry* **2013,** *113*, 96-101.

10. (a) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E., *Phys. Rev. Lett.* **2003,** *91*, 146401; (b) Schwabe, T.; Grimme, S., *Phys. Chem. Chem. Phys.* **2006,** *8*, 4398-4401.

11. Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K., *J. Chem. Phys.* **2001,** *115*, 3540-3544.

12. Ruggiero, M. T.; Gooch, J.; Zubieta, J.; Korter, T. M., *J. Phys. Chem. A* **2016,** *120*, 939-947.

13. Grimme, S., *J. Chem. Phys.* **2006,** *124*, 1-16.

14. (a) Chai, J. D.; Head-Gordon, M., *J. Chem. Phys.* **2008,** *128*, 1-16; (b) Staroverov, V. N.; Scuseria, G. E.; Tao, J. M.; Perdew, J. P., *J. Chem. Phys.* **2003,** *119*, 12129-12137; (c) Becke, A. D., *J. Chem. Phys.* **1993,** *98*, 5648-5652.

CHAPTER 6: PUTATIVE METALLOPROTEIN FROM CLOSTRIDIUM DIFFICILE MAY BE INVOLVED IN SPORE COAT FORMATION

SECTION 6.1. INTRODUCTION

***Clostridium difficile* infections and treatment options.** *C. difficile* is a nosocomial pathogen responsible for infections such as antibiotic-related colitis that, in the case of immunocompromised patients, may result in patient death.[1] It is a gram-positive bacillus bacteria. The effectivity of transmission of the bacteria can be attributed the formation of a robust endospore.[1] The general spore morphology of *Clostridium difficile* pathogenic strain 630 (CD630) was identified through transmission electron microscopy (TEM) (**Fig. 6.1**).

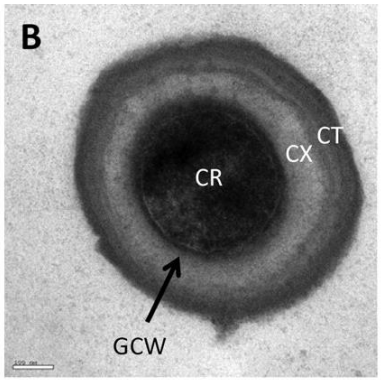
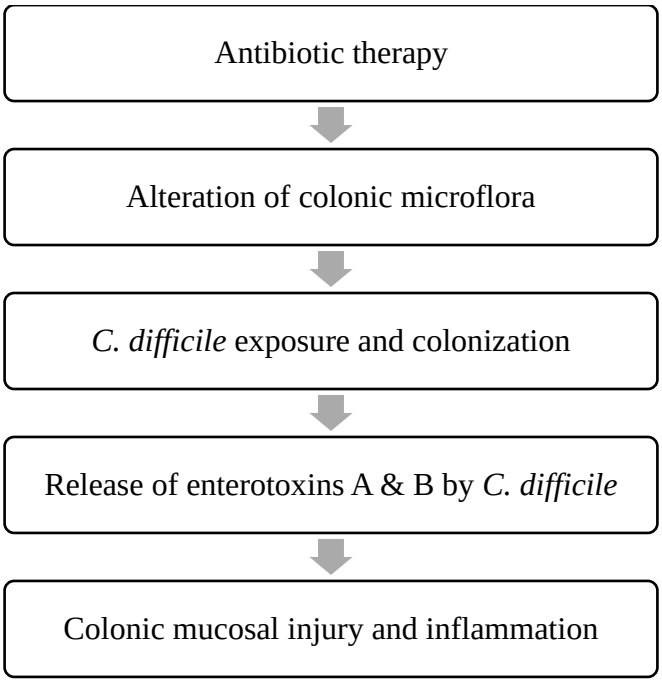


Figure 6.1. TEM image of the general structure of a mature *Clostridium difficile* 630 spore. Abbreviations are as follows: CR, core; GCW, germ cell wall; CX, cortex; CT, coat. Figure taken from reference 12.

The formation of a spore coat allows the bacteria to survive in unfavorable conditions (any environment outside a human host).[2] Hospital patients are likely to incur a *C. difficile* infection during their stay due to the nature of pathogenesis (**Scheme 6.1**).



Scheme 6.1. Pathogenesis of *C. difficile* infection. Scheme adapted from reference 1.

The first step of pathogenesis leads to the disruption of the natural microflora between in the patient's colon; this is a side effect of common antibiotics such as ampicillin, amoxicillin, and clindamycin.[1] In healthy patients, the natural colonic microflora is usually capable of preventing the colonization of *C. difficile*.[3] However, exposure to the *C. difficile* organism or dormant spore can quickly lead to an infection when this natural defense is weakened or absent. Once *C. difficile* is able to germinate in the colon, enterotoxins A and B are released and damage the epithelial cells of the colon resulting in inflammation and eventually mucosal damage.[4] After germination, more *C. difficile* endospores are produced (by sporulation) resulting in rapid transmission from infected patients or healthcare workers to uncolonized hosts.[5] Excessive production of the enterotoxins may lead to severe

(or fulminant) colitis; symptoms include megacolon (severe inflammation) and perforation which, when left untreated, can lead to death.[6]

Treatment for *Clostridium difficile* infections. *C. difficile* would be unable to proliferate let alone survive outside the human host without the metabolically dormant spore state.[7] For mild cases, medical intervention is not required and patients often recover from this infection after their natural microflora repopulates the colonic tract.[8] Allowing the normal colonic microflora to reconstitute without additional antibiotics greatly decreases the likelihood of a repeat infection. In more aggressive cases, a course of vancomycin or metronidazole is the recommended course of action. Vancomycin is the more effective antibiotic; but generally not the first choice for treatment because not only is it more expensive, but also the overuse of vancomycin can lead to the proliferation of a vancomycin-resistant strain of bacteria.[1] However, a round of vancomycin will only reduce the population of vegetative cells of *C. difficile* but not the dormant (endospore protected) cells.[9] Aggressive cleaning methods such as high concentrations of hydrogen peroxide and chlorine-containing disinfectants can only remove sporulated cells from infected surfaces.[10]

Due to the nature of pathogenesis and current treatment options, infected patients are predisposed to relapse again.[10] Thus, a new treatment that can target *C. difficile* infections without disrupting the normal gut microbiome of the patient is desperately needed as more antibiotic resistant strains of *C. difficile* are being discovered.

Targeting endospore forming proteins as an alternative treatment pathway for *C. difficile* infections. The dormant spore state of *C. difficile* is the transmissible and infectious morphotype. The metabolic dormant nature of the endospore form makes this bacteria largely antibiotic resistant.[9] Mutants of *C. difficile* that do not form the robust endospore are unable to transmit the disease.[11] The nature of spore formation is not well understood and very little is known about the composition or method of formation of the *C. difficile* spore coat. Part of the mystery of the spore formation in *C. difficile* arises from the unique proteome; there are few similarities between this proteome and that of the well-studied *Bacillus* genus.[5]

To resume vegetative growth and proliferation inside the human host, *C. difficile* cell must shed the endospore coat. This makes the cell vulnerable to antibiotics and other treatments (lysozyme, ethanol, and heat).[7] By better understanding the components of the spore formation and maturation in *C. difficile*, a new treatment could be developed that targets the spore form. This chapter will discuss the spectroscopic characterization of the protein CotA which would have implications in the role it may play in spore formation.

The coat (Cot) family of proteins are associated with the endospore of *C. difficile*. A preliminary study in 2011 identified several proteins related to the spore coat formation in *C. difficile*.[12] Through bioinformatics analyses, five proteins were identified as coat (Cot) proteins that express on the outer coat layers of CD630: CotA, CotB, CotCB, CotD, and CotE. It was proposed that the Cot proteins were involved in the polymerization and maturation of the endospore coat.[12] Their specific roles were later narrowed down to different spore structures: CotA was associated with the outer layer of the spore coat, CotB

was associated with the exosporial layer, while CotCB, CotD, and CotE were associated with both layers.[7] The most promising Cot protein target was CotA; several publications from the Cutting group reported the malformation of the spore coat in *cotA* knockout variant (**Fig. 6.2**)

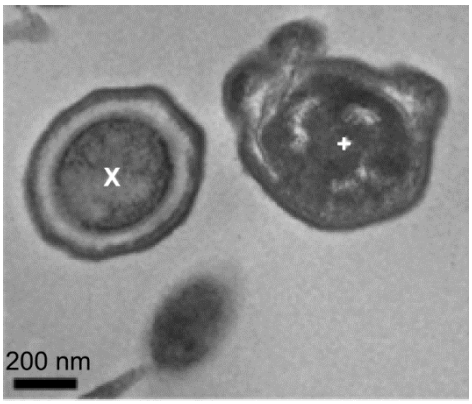


Figure 6.2. Cells exhibiting normal spore morphology (left, "X") and the *cotA*$^{-}$ mutant phenotype (right, "+"). Figure taken from reference 7.

A preliminary UV-vis (Abs) absorption spectrum suggested an iron-sulfur cluster (ISC) was present in native CotA; the intensity of the feature at 407 nm aligned with that of a charge transfer (CT) band which would be explained by sulfur-to-iron charge transfer (a ligand-to-metal charge transfer, LMCT) (**Fig. 6.3**).

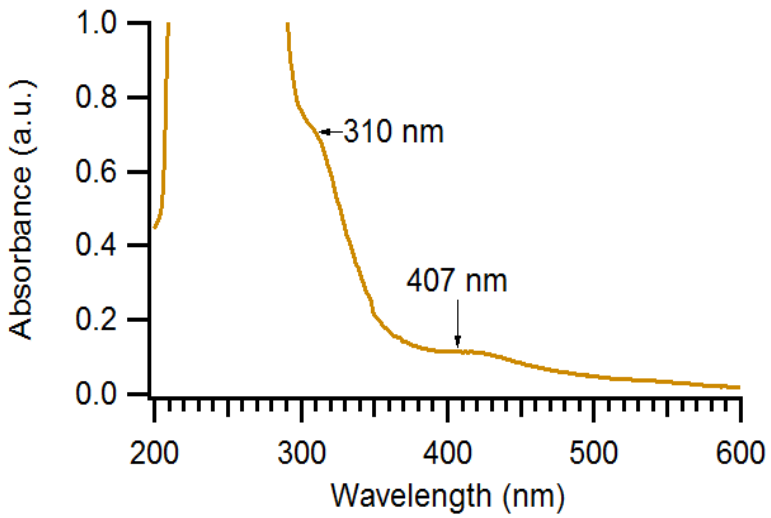


Figure 6.3. Abs spectrum of CotA in 625 mM KPi and 1250 mM NaCl buffer at pH 7.4.

Bioinformatics used to identify potential ISC binding sites in CotA. The sequence of CotA has no shared homology with other proteins in existing databases.[12] Since the CotA sequence bears no resemblance to the proteins of the protein data bank, the program MetalDetector v 2.0 was used to better understand the sequence of amino acids. MetalDetector v2.0 analyzes the sequence for cysteine and histidine residues that could be involved in transition metal binding sites.[13] MetalDetector v2.0 assesses the primary structure of a protein and predict transition metal binding sites, thus MetalDetector allows us to better understand the CotA active site without any structural information.[13] Comparing the CotA sequence could help identify which iron-sulfur clusters are present. There are 15 cysteine residues (C, highlighted) in the CotA sequence that could be ligands in the ISC metal site (**Fig. 6.4**).

*MGHHHHHHDYDIPTTENLYFQ*GAHMGIQRPTSTVENNKCREDFRFTQEYE
EDYPNTNERYYENYQVADRYYNYPNKYKEPKIKQCCCKKSMREALELLRY
DALRPFVNFNQFAFISDFFIVGANLVGIDLSAPPKDNLSGLDGTFERFSACN
CDLIDIAGRVSYPIPVPLTLEGLINTIGTIPGVAELIALIDAVIPPTIDLGAILDAIL
AAIIDFILAASTPLANVDLASLCNLKAVAFDITPADYEDFIASLGYYLDKKHYK
ECNCNCDCDDCCCNKGILDNLYMSNINNQVTVVAGSLVLTGVEVLGKKND
VIVLGNSNDSRIYFVCVDSIDYIA

Figure 6.4. The CotA sequence with the N-terminal His_6-tag. Italicized residues are cleaved by TEV Protease. Cysteine residues are highlighted and histidine residues are in blue text.

Other amino acids, such as histidine (H, blue text) could bind the iron atom in the Rieske cluster.[14,15] The software program MetalDetector v2.0 predicted two metal binding sites highlighted in the sequence in magenta (site 1) and green (site 2). Neither site 1 nor site 2 matches any of the common motifs of [2Fe-2S], Rieske [2Fe-2S], [3Fe-4S], [4Fe-4S] or HiPIP [4Fe-4S] clusters.[14] There is a [4Fe-4S] motif found at the C-terminus of the sequence, CX_3CX_2C (site 1), but it has been reported that this motif only forms ISC if located at the N-terminus of a protein sequence.[16] However, there are not consensus motifs for ISC clusters throughout nature so clusters must be characterized through spectroscopy.

The lack of common FeS cluster binding motif is not surprising as the CotA sequence did not have any shared homology with published sequences of other metalloproteins.[12] This may indicate that the ISC conformation has not been published or that the cluster is formed from a multimeric CotA species. If the cluster is formed through the multimerization process, then a previously published binding motif would not be found

in the CotA sequence. Either way, the lack of ISC motif likely signifies that the motif binding the ISC is rare or that it has not been previously reported.

In a collaboration with Professor Aimee Shen (University of Vermont, Department of Microbiology and Molecular Genetics currently Sackler School of Graduate Biomedical Sciences at Tuft University, Boston, MA), the protein CotA was been investigated. The putative metal-centric active sites were probed spectroscopically with the goal of understanding how metal binding in these proteins is related to their involvement in spore formation.

SECTION 6.2. BIOLOGICAL METHODS, RESULTS AND DISCUSSION

Expression of CotA in the pUC57 plasmid. Initially, the Shen lab cloned the *cotA* gene into a pUC57 plasmid encoded with a C-terminal cysteine protease domain (CPD) with a His_6-tag.[17] This fusion protein tag is used to increase solubility of the fused proteins to increase the yield of the protein of interest. A phytic acid cleavage site (VDAL/ADGK) was engineered into the pUC57 plasmid to separate the dual protein construct into the CPD tag and the CotA protein.[17] The protein construct must be cleaved to ensure any FeS features present in subsequent spectroscopic investigations arise solely from the ISC of CotA. The free His_6-tag could bind to free metals such as iron; this would result in inaccurate characterization of the ISC in CotA.

Efforts to cleave CotA from the CPD fusion tag were made in parallel with researchers of the Shen lab but were ultimately unsuccessful. The failed cleavage may be because the protein construct was folded so that the phytic acid binding site was inaccessible. To ensure cleavage of CotA from a His_6-tagged construct, the gene was inserted into a vector containing a tobacco etch virus (TEV) cleavage site.

Expression of CotA in the pET15b plasmid. The *cotA* gene was cloned from the pUC57 plasmid into a modified pET15b plasmid (gift from Eric Skaar, Vanderbilt University) equipped with the TEV cleavage site (ENLYFQ/G) with the assistance of Kristin Schutz (Shen Lab). A pair of restriction enzymes, SalI and BamHI, were used to cut the *cotA* gene from the pUC57 plasmid and insert the gene into pET15b (**Fig. S1**). The insertion of *cotA* into pET15b was confirmed by DNA sequencing provided by the University of Vermont Cancer Center (VCC, University of Vermont College of Medicine).

The stock was transformed from a DH5α cell line (used for cloning) into the BL21-GOLD (DE3) cell line (suitable for high-level expression).

Purification of CotA from *Escherichia coli*. *E. coli* is used as the expression host to produce the protein of interest, CotA. The plasmid containing the protein of interest functions in the host cell such that, when induced, will lead to high levels of gene expression. High expression is required to produce many copies of CotA to be used for spectroscopic investigation. The ideal conditions for optimal CotA expression in *E. coli* were identified through rigorous testing (**Table 6.1**).

Table 6.1. Growth parameters altered to optimize the expression of CotA in BL21-GOLD (DE3) cell line in *E. coli*.

Media	Induction	Growth time (Temperature)	Was chelated iron cocktail[c] added?
LB[a]	IPTG	4 hr (30 °C)	No
LB[a]	IPTG	18 hr (30 °C)	No
LB[a]	IPTG	8 hr (30 °C)	No
TB[b]	Autoinduction	26 hr (30 °C)	Yes
TB[b] (no glycerol)	Autoinduction	48 hr (30 °C)	Yes
TB[b]	IPTG	4 hr (30 °C)	Yes
TB[b]	IPTG	4 hr (37 °C) 18 °C after induction	Yes
TB[b]	Autoinduction	72 hr (20 °C)	Yes

[a] Lysogeny broth (10 g NaCl, 10 g tryptone, 5 g yeast, 1 mL ampicillin per 1 L)
[b] Terrific broth (24 g yeast, 12 g tryptone, 4 mL glycerol, 100 mL 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4 per 1 L)
[c] Sterile filtered 22 mM $FeSO_4 \cdot 7H_2O$, 1.6 mM EDTA in MilliQ water added at time of induction

The pET15b plasmid contains the *lacI* operon thus induction by isopropyl β-D-1-thiogalactopyranoside (IPTG) and autoinduction are both viable methods of CotA expression. In order to collect the protein of interest, frozen cell pellets were resuspended in 500 mM NaCl, 50 mM Tris, 15 mM imidazole, 2 mM DTT, buffer at pH 7.4 prior to lysis by sonication. Sonication uses cavitation to burst open the cells where air bubbles are formed and burst thus producing small shock waves throughout the resuspended cell liquid. These shock waves lyse the bacteria and breakdown DNA. Pelleted cells were slowly brought to room temperature then resuspended in the aforementioned lysis buffer. The cells were lysed at 60 % sonication duty for six cycles of one minute of sonication followed by a one min rest period. A sonifier equipped 13 mm tapped step horn tip on a threaded body was used to lyse the *E. coli* cells. To minimize protein degradation, the lysate was kept on ice during sonication. In order to separate the insoluble debris, the sample was centrifuged at 15,000 x*g* at 4 °C for 30 to 60 min and the supernatant was further filtered through a 0.45 μm membrane (Millipore).

The CotA monomer (approximately 35 kDa) as well as higher order multimers were observed in the sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel (**Fig. 6.5**). Autoinduction was shown to have superior expression of CotA dimer and monomer. Expression was ultimately optimized by growing the BL21-GOLD (DE3) *E. coli* strain at 20 °C in Terrific Broth (a more nutrient-rich alternative to Luria-Bertani media) with a supplemental iron source. Growing at this low temperature enhances solubility which can directly increase protein yield.[18] As such, subsequent expression attempts of CotA followed these parameters.

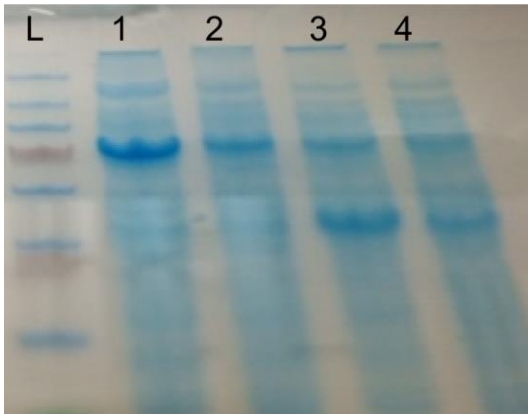


Figure 6.5. Comparing efficacy of methods of expression in pET15b through SDS-PAGE gel. The lanes represent: (L) molecular weight ladder; (1) lysate of autoinduced cells; (2) lysate of IPTG induced cells; (3) filtered lysate of autoinduced cells; and (4) filtered lysate of IPTG induced cells. Frozen cell pellets were resuspended in 500 mM NaCl, 50 mM Tris, 15 mM imidazole, 2 mM DTT, 5 μM buffer at pH 7.4 prior to lysis.

The filtered lysate was then loaded onto a HisPur Ni-NTA column (Pierce) previously equilibrated with 500 mM NaCl, 50 mM Tris, 2 mM DTT buffer at pH 7.4. The Ni-NTA column was then washed with: 500 mM NaCl, 50 mM Tris, 2 mM DTT buffer at pH 7.4; 500 mM NaCl, 50 mM Tris, 2 mM DTT, 25 mM imidazole buffer at pH 7.4; 500 mM NaCl, 50 mM Tris, 2 mM DTT 50 mM imidazole buffer at pH 7.4; 500 mM NaCl, 50 mM Tris, 2 mM DTT, 100 mM imidazole buffer at pH 7.4; and 500 mM NaCl, 50 mM Tris, 2 mM DTT, 200 mM imidazole buffer at pH 7.4, 500 mM NaCl, 50 mM Tris, 2 mM DTT, 400 mM imidazole buffer at pH 7.4. Solely multimeric CotA eluted in the 500 mM NaCl, 50 mM Tris, 2 mM DTT, 200 mM imidazole buffer at pH 7.4 (**Fig. 6.6**).

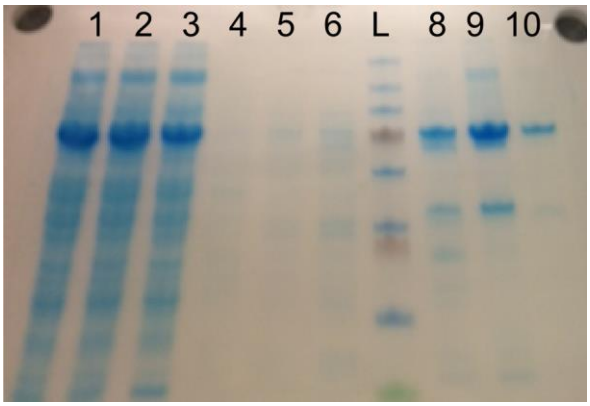


Figure 6.6. Purification of CotA as monitored by SDS-PAGE. The lanes represent: (1) unfiltered lysate; (2) filtered lysate; (3) flow through; (4) 500 mM NaCl, 50 mM Tris, 2 mM DTT buffer at pH 7.4 (equilibration buffer); (5) 25 mM imidazole in equilibration buffer; (6) 50 mM imidazole in equilibration buffer; (L) molecular weight ladder; (8) 100 mM imidazole in equilibration buffer; (9) 200 mM imidazole in equilibration buffer; and (10) 400 mM imidazole in equilibration buffer.

The elution of CotA in a wash containing a higher concentration of imidazole indicates that the protein is bound at Ni(II)-NTA through the His_6-tag. It was proposed that the active form of CotA may be multimeric *in vivo* and was predominantly observed to be in the dimeric form during sample preparation.

Removal of His_6-tag. Once the procedure for purifying CotA had been optimized, efforts were focused on the cleavage reaction with TEV to remove the His_6-tag. Histidine-tags can bind to available metals and a supplemental iron source was added during induction. TEV was used to cleave the His_6-tag from CotA. The S219V TEV cell line was a gift from David Waugh (National Cancer Institute). TEV was expressed and purified as previously reported.[19] Purified TEV was added to CotA in an OD_{280} 1:100 ratio in 500 mM

NaCl, 200 mM imidazole, 50 mM Tris, 2 mM DTT, 0.5 mM EDTA buffer at pH 7.4 and allowed to cleave overnight at 4 °C. Cleaved protein was then dialyzed against equilibration buffer to decrease imidazole content (to prevent unwanted elution of tagged protein) and to remove cleaved His_6-tag.

Tag-free CotA was separated from the mixture by Ni(II)-NTA affinity chromatography; without the His_6-tag, CotA is collected in the column flow through. Prior to loading, the Ni-NTA column was equilibrated with 500 mM NaCl, 50 mM imidazole, 50 mM Tris, 2 mM DTT buffer at pH 7.4. The cleaved protein was concentrated using in an Amicon stirred cell (Millipore) at low pressure (20 psi compared to the optimal 55 psi) due to the labile nature of the FeS cluster. The concentration of protein (all multimers of CotA) was determined using the colormetric Bradford assay (Pierce) using bovine serum albumin as a protein standard. For spectroscopic investigation, CotA was exchanged into 625 mM potassium phosphate, 1250 mM NaCl, 2 mM DTT buffer at pH 7.4 using a PD-10 desalting column (GE Healthcare). The protein was then mixed with glycerol in a 40:60 ratio (*v/v* %) to form a transparent glass upon freezing for MCD spectroscopy.

Oxidation leads to FeS cluster degradation. The absorption spectrum of purified CotA had a characteristic charge transfer band at 415 nm and unresolved features between 500-600 nm were also observed. The color of the eluent, a straw-like color, supported the proposition that CotA was a metalloprotein and most likely an ISC protein.[20] Ambient oxidation appeared to be linked to a loss of cluster (as indicated by Abs spectroscopy); ISCs are often air-sensitive and oxidation may lead to cluster degradation.[21]

The addition of dithiothreitol, DTT, slowed the loss of the cluster as monitored by Abs spectroscopy due to its reducing properties (E = -330 mV) which likely combats the ambient oxidation by molecular oxygen. Excess DTT will not reduce disulfide bonds into two thiol cysteines if they are inaccessible to the solvent; excess DTT could lead to partial or complete unfolding of the protein. Multiple cluster configurations could arise from the degradation from ambient oxidation; the oxidation of a [4Fe-4S] cluster usually leads to the loss of an iron and the irreversible formation of a [3Fe-4S] or [2Fe-2S] cluster or may lead to compete cluster loss.[22, 16] With excess reductant, such as sodium dithionite, [3Fe-4S] and [2Fe-2S] clusters have been shown to combine in such a manner to form a new [4Fe-4S] cluster.[16] While DTT was required to lengthen the lifetime of FeS cluster bound CotA, this may have caused a non-native ISC to form though the combination of multiple clusters or ambient oxidation could have led to partial degradation of the cluster to a semi-stable ISC.

SECTION 6.3. SPECTROSCOPIC METHODS, RESULTS, AND DISCUSSION

Initial investigation indicated presence of ISC in CotA. A preliminary Abs absorption spectrum implicated an iron-sulfur cluster (ISC) was present in native CotA. This spectrum was collected using a Cary 100 Bio Spectrophotometer. Data was collected between 800 and 200 nm using a scan rate of 600 nm/min, a data interval of 0.5 nm, and an integration time of 0.1 s. The intensity of the feature at 407 nm aligned with that of a charge transfer (CT) which would be explained by sulfur-to-iron charge transfer (a ligand-to-metal charge transfer, LMCT) (**Fig. 6.4**).

CT features found in the energy range (22,000 to 25,000 cm^{-1}) have been documented in the literature as a sulfur-to-iron charge transfer for many ISC metalloproteins such as [2Fe-2S], Rieske [2Fe-2S], pseudo-cubane [3Fe-4S], as well as [4Fe-4S] clusters.[14] The near-UV feature (310 nm) is characteristic of rubredoxin, [2Fe-2S], and Rieske [2Fe-2S] clusters.[14] The chromophore responsible for the intense UV feature is the protein (specifically the peptide backbone as well as aromatic amino acids: phenylalanine, tryptophan, and tyrosine). As observed in **Figure 6.3**, many overlapping CT bands lead to the broadening of the absorption spectra of ISC proteins; there are unresolved features responsible for the broad weak peak from 450-600 nm.

While absorption spectroscopy is well-suited for identifying the presence of an ISC, it cannot be used to characterize the cluster alone. Due to the nature of the technique, many overlapping bands may be present in an Abs spectrum that lead to the broadening of features.[23] Other methods, such as magnetic circular dichroism (MCD) spectroscopy, would be required in order to quantify the number of ISCs present, the type of ISC, as well as their respective oxidation states. Because MCD's selection rules are favorable to detect FeS features and band resolution is improved at cryogenic temperatures, a spectral library of FeS cluster conformations and oxidation states have been cataloged in the literature using MCD spectroscopy.[23a, 24]

Characterizing iron-sulfur clusters with spectroscopic techniques. Many spectroscopic techniques have been used to identify and characterize FeS clusters of a metalloprotein of interest. The variety of useful techniques stems from the inherent magnetism and electron-rich structure of the many ISC. The [2Fe-2S] and cubane [4Fe-4S] clusters are the most prevalent ISCs.[25] There are many variations on the FeS cluster motif and each structure has many accessible oxidation states and spin states (**Fig. 6.7**).

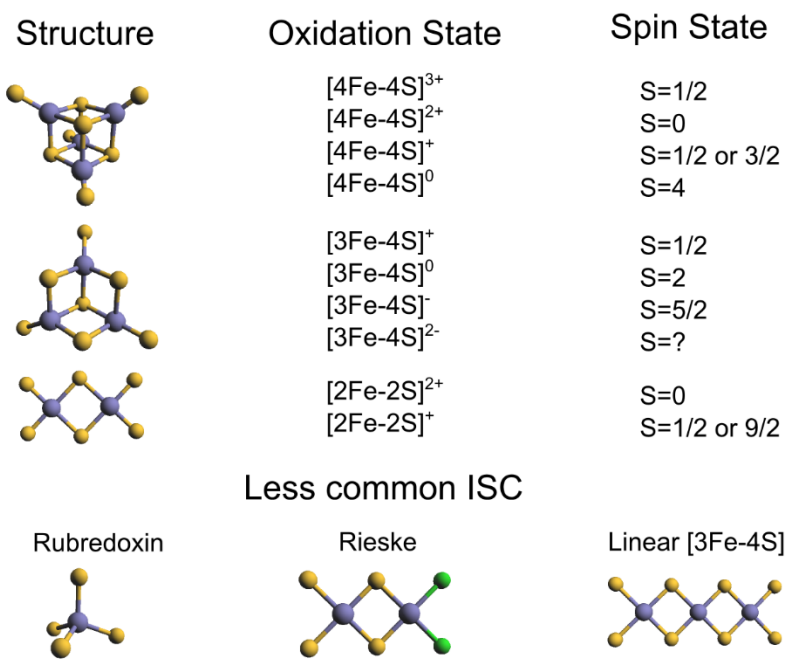


Figure 6.7. Some iron-sulfur clusters with their potential oxidations states and spin states. The structure of [4Fe-4S] cluster is cubane, the structure of the [3Fe-4S] cluster is pseudo cubane, and the [2Fe-2S] is a linear cluster. Silver spheres represent iron (accessible oxidation states include Fe^{2+}, $Fe^{2.5+}$, and Fe^{3+}). Golden spheres represent sulfur (cysteine or sulfide). Green spheres represent non-sulfur ligands (such as nitrogen from histidine) in a Rieske cluster.

Two methodologies used to probe metalloproteins containing ISC successfully are Mössbauer spectroscopy and electron paramagnetic resonance (EPR) spectroscopy.[26] EPR spectroscopy can provide information pertaining to the structure and function of many types of ISC proteins as was the first technique used to investigate ISC proteins.[27] This technique can be used to provide the following pieces of information pertaining to the ISC present in a metalloprotein: redox states, electronic structure, as well as proximity of many

clusters (if multiple clusters are present).[27] EPR spectroscopy can characterize the spin state of the cluster; in the typical perpendicular configuration, a spectrum without resonances would indicate an integer-spin species.[23a] Mössbauer spectroscopy is inherently specific to the FeS cluster ISC proteins because the technique relies of the Fe-57 isotope (a naturally abundant isotope, 2.1%).[25] This technique provides information on the oxidation and spin state of the iron atom(s) of the cluster.[23a]

Multiple techniques are often required to be used in tandem to characterize an ISC protein. Other methods used to study FeS proteins include electron–nuclear double resonance (ENDOR) spectroscopy, electron spin echo envelope modulation (ESEEM) spectroscopy, X-ray absorption spectroscopy, circular dichroism (CD) spectroscopy, and magnetic circular dichroism (MCD) spectroscopy. The simplest experiment to collect information pertaining to the presence of the ISC is Abs spectroscopy as CT features (sulfur-to-iron) are easily observed, $1 \leq \varepsilon \leq 100$ M^{-1} cm^{-1}. [23a]

MCD will be used to identify the ISC in CotA. MCD spectroscopy can be used to identify the metal center of a metalloprotein.[23a] If the metalloprotein has an ISC, MCD can be used to differentiate between possible FeS cluster configurations, oxidation states, and spin states.[24a] The diagnostic nature of MCD spectra for FeS clusters allows this technique to be used as a fingerprinting tool to identify an unknown ISC in a protein. MCD spectroscopy provides a wealth of information when paired with Abs and CD spectroscopy. These three techniques (Abs, CD, and MCD spectroscopies) probe the same electronic transitions but differ in their selection rules so these transitions will have different intensities in each spectroscopic technique (Equations 1-3).

$$I(Abs) = \langle \Psi_i | \hat{\mu}_{el} | \Psi_f \rangle \quad \textbf{Eqn 1.}$$

$$I(CD) = \langle \Psi_g | \hat{\mu}_{el} | \Psi_i \rangle \langle \Psi_g | \hat{\mu}_{mag} | \Psi_e \rangle \quad \textbf{Eqn 2.}$$

$$I(MCD) = \gamma \beta_B B \left[\left(-\frac{f(E)}{E} \right) A_1 + \left(B_0 + \frac{C_0}{kT} \right) f(E) \right] \quad \textbf{Eqn 3.}$$

MCD provides information on the ground state and excited state. For transition metal complexes, the energies of the d-orbitals become non-degenerate due to the magnetic field and cryogenic temperatures influence the population of the d-orbitals. The three types of MCD features (A-term, B-term, and C-term) each have unique characteristics that make them identifiable. A-term features have a unique derivative shape in MCD spectra; these transitions arise from a degenerate excited state and are temperature independent. B-term features are also temperature independent but arise from the mixing of near-degenerate states induced by the magnetic field. C-term features are the only temperature dependent feature in an MCD spectra; they arise from the degenerate ground state. The C-term is temperature dependent because the Boltzmann distribution dictates the population of these degenerate ground states.

The MCD spectra are of a 16 μM CotA sample in 60 % glycerol and 40 % 250 mM potassium phosphate, 500 mM NaCl, 5 mM DTT buffer at pH 7.4. This sample was loaded into a home-made copper and quartz cell which was flash-frozen in liquid nitrogen. Spectra were collected with a home-built MCD setup composed of a Jasco J815 spectropolarimeter and an Oxford SM400-8T Spectromag which was controlled by a Mercury iTC temperature controller and powered by the Mercury iPS power supply. Samples were scanned from 800

to 350 nm at a speed of 200 nm/min, a pitch of 0.5 nm, an integration time of 0.25 sec, and a bandwidth of 1 nm. For these MCD spectra, the natural CD signal was removed by subtracting the negative field trace from the positive field trace then divide by two. The near-IR region (λ= 700-1100 nm) was also scanned, but no features were detected over the noise of the baseline. At wavelengths lower than 300 nm, protein-based features will dominate the spectra. At wavelengths lower than 360 nm, excess dithionite will cause spectral interference. For these two reasons, 350 nm was selected as the cutoff for the MCD spectra. The noise in these spectra arose from two sources: mild frosting of the quartz discs of the sample cell and a low concentration of holo-protein relative to apo-protein. The arbitrary intensity unit, millidegrees or "mdeg", was used because an accurate $\Delta\varepsilon$ could not be obtained due to an unknown mixture of species.

First, we will discuss the temperature dependent MCD spectra of CotA (**Fig. 6.8**). The temperature-dependent MCD spectra show C-term features at 19,100 and 24,000 cm^{-1}; the trace collected at the lowest temperature, 5 K, has the most intense features. This is indicative of a paramagnetic species which can help eliminate many ISC as the potential cofactor of CotA. The diamagnetic ISC that are unlikely to be present in CotA include $[2Fe\text{-}2S]^{2+}$, Rieske $[2Fe\text{-}2S]^{2+}$, $[2Fe\text{-}2S]^{0}$, $[3Fe\text{-}4S]^{0}$, $[4Fe\text{-}4S]^{2+}$ and $[4Fe\text{-}4S]^{0}$; the antiferromagnetic coupling of multiple iron centers leads to S_{total}=0.

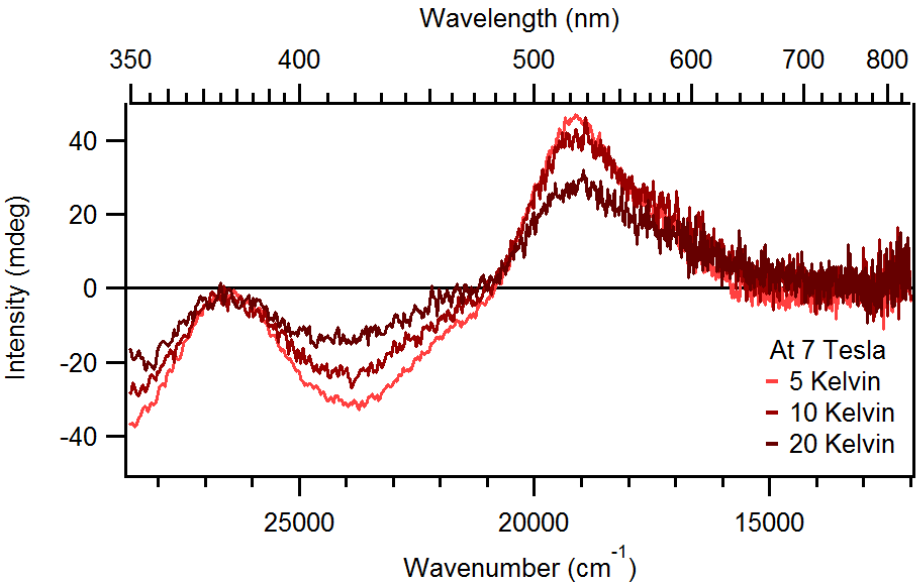


Figure 6.8. Temperature-dependent MCD spectra of CotA in 60 % glycerol and 40 % 250 mM potassium phosphate, 500 mM NaCl, 5 mM DTT buffer at pH 7.4. at 7 Tesla. Darker color traces indicate warmer temperatures.

The field-dependent MCD spectra of CotA is indicative of a paramagnetic species (**Fig. 6.9**). The intensity of the features saturated at a magnetic field above 3 Tesla. This saturation is characteristic of a high-spin species. The MCD features should allow for the identification of the oxidation state as well as cluster conformation of the bound FeS cluster (**Table 6.2**).

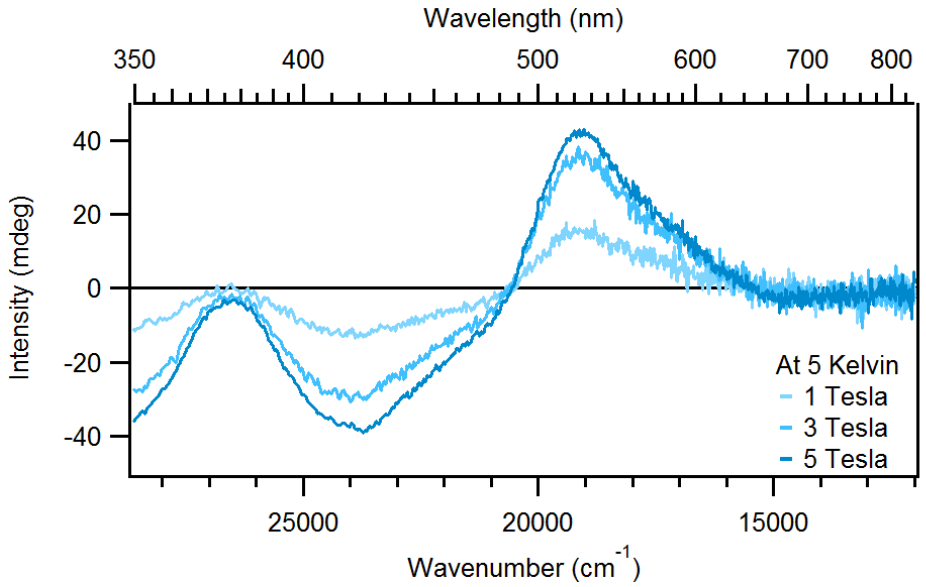


Figure 6.9. Field-dependent MCD spectra of CotA in 60 % glycerol and 40 % 250 mM potassium phosphate, 500 mM NaCl, 5 mM DTT buffer at pH 7.4. at 5 K. The darker colored traces indicate higher magnetic fields.

Table 6.2. Summary of MCD features of CotA

Feature	λ (nm)	ν (cm^{-1})
Zero point	377	26600
Minima	418	24000
Zero crossing	477	21000
Maxima	522	19100
Zero point	633	15300

The MCD spectra of the visible region of CotA is unlike the previously published spectra of ISC-containing proteins (**Figs. 6.8** and **6.9**). This could be because there is a mixture of FeS clusters in the sample, CotA is a new type of metalloprotein, or CotA contains a novel ISC. There are two temperature-dependent features present in the MCD spectra. This temperature dependence implies that the chromophore is a paramagnetic

species.[23a] The minimal field dependence of the features indicates that the features arises from a high-spin (S>1/2) species (**Fig. 6.9**). There are many examples in the literature of high-spin ISC of various configurations.[28,29] An exact spectral match was not expected because there is a mixture of multimers in the sample and multiple oxidation states are also possible. Other MCD-active protein cofactors and substrates will be considered as well to further support the claim that CotA is binding one (or more) iron-sulfur clusters.

MCD spectra of CotA does not match spectra of ISC in the literature. Several oxidation states can be eliminated as the ISC bound by CotA due the way the MCD sample was prepared. Firstly, the addition of DTT lowered the reduction potential of the solution to -330 mV. Protein with a reduction potential less than -330 mV are more likely to be oxidized while protein centers with reduction potentials greater than -330 mV have a greater likelihood of being reduced. Each cluster type has a range of reduction potentials and the oxidized and reduced cluster both are accessible at potential of -330 mV for most ISC proteins (**Fig. 6.10**).

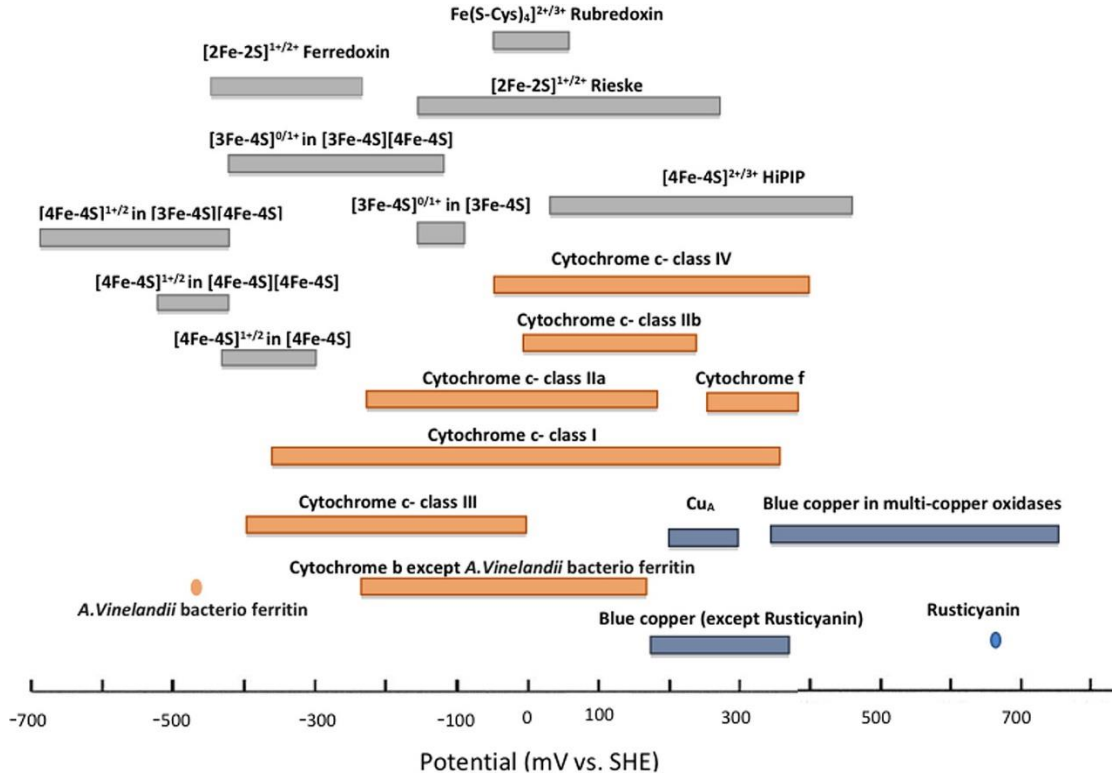


Figure 6.10. Reduction potentials of common redox centers. Iron-sulfur proteins are indicated in gray, heme proteins are indicated in orange, and copper proteins are shown in blue. Taken from reference 14.

Due to their reduction potentials, Rieske [2Fe-2S] and [3Fe-4S] clusters are likely present in their reduced forms.[14] Thus [3Fe-4S] clusters are eliminated from the list of likely ISC of CotA because the reduced form, $[3Fe\text{-}4S]^0$, is a diamagnetic species which is in disagreement with the MCD spectra. Other iron-sulfur clusters were determined to be unlikely ISC for CotA based on discrepancies between their Abs and MCD spectra and the spectra of CotA. Based on the reduction potential, any Rubredoxin 1Fe cluster would be in the reduced state; 1Fe(II) would have an MCD spectra with an intense derivative-shaped peak centered at 325 nm where all features are temperature dependent.[30] CotA most likely

does not have a Rubredoxin-like cluster. Based on their reduction potential, [2Fe-2S] clusters could be either oxidized or reduced (**Fig. 6.10**); to further complicate things, there are many accessible spin states of [2Fe-2S] clusters.[28] However, neither the Abs or MCD spectra of these [2Fe-2S] clusters are reminiscent of the spectra of CotA; therefore, a [2Fe-2S] cluster is unlikely to be present in CotA.[14, 24a]

A Rieske [2Fe-2S] cluster is likely ISC bound by CotA. The major spectra features (negative feature at 418 nm and the positive feature at 520 nm) have been tentatively assigned to arise from a reduced Rieske [2Fe-2S] cluster. Rieske [2Fe-2S] clusters are bound to the protein through amino acids other than cysteines.[14] Ligating groups that could bind iron include imidazole, phenolate, carboxylate, or hydroxyl groups; the two non-cysteine ligands can be from the same type of amino acid or could be different amino acids.[15] Rieske clusters have MCD features at (positive feature) 380-383 nm, (negative feature) 425-432 nm, and (positive feature) 505-550 nm.[14,15]

Instead of a positive feature at approximately 380 nm, there is a zero point feature in the CotA MCD spectra. The absence of the 380 nm feature is likely due to overlapping features arising from DTT as well as the protein. A low temperature CD spectrum of CotA could aid in definitively assigning the ISC present in CotA as a reduced Rieske center; the CD spectrum of a reduced Rieske protein would have an intense negative feature at 384 nm, two (or more) bands between 400 and 480 nm, negative band at 500 nm, and a weak feature at 660 nm.[15] The CD spectrum collected at 5 Kelvin is limited to a single distinct UV feature is composed of many overlapping bands arising from, but not limited to, DTT and aromatic amino acids (**Fig. 6.11**).

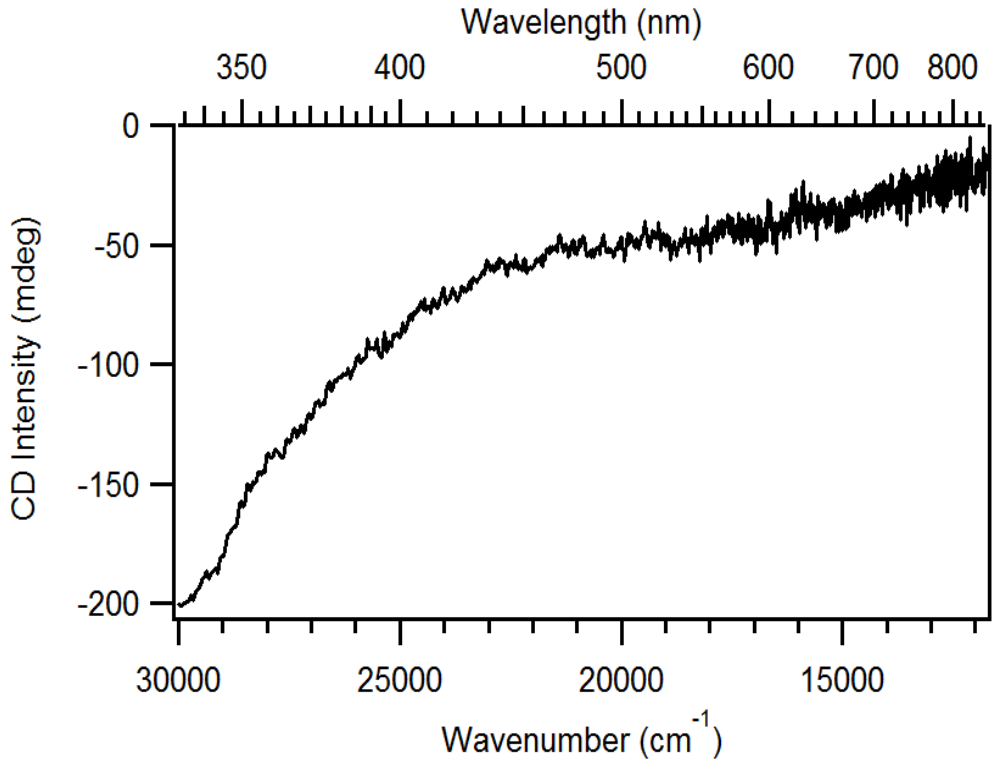


Figure 6.11. CD spectrum of CotA at 5 Kelvin

The intense negative MCD feature at 384 nm is likely hidden due to the many overlapping bands responsible for the convoluted spectrum. Any bands arising from the apo-protein present in the sample would overpower less intense features; because of significant cluster degradation, there is likely an excess of apo-protein as compared to holo-protein. While these Rieske bands could not be discerned from the convoluted spectrum, the MCD spectra still support a reduced Rieske ISC is present in CotA.

A S=3/2 [4Fe-4S]$^{+}$ cluster is likely ISC bound by CotA. There could be other ISC configurations contributing to the CotA MCD spectra. The intense positive feature at approximately 520 nm is a common feature in [4Fe-4S]$^{+}$ clusters for both S=1/2 and S=3/2.[29] While the overall charge of the [4Fe-4S] cluster could be either 1+ or 2+ due to

the reduction potential of DTT, only $[4Fe\text{-}4S]^{+}$ would contribute to the C-term MCD spectra (**Fig. 7**). This is because $[4Fe\text{-}4S]^{2+}$ is a diamagnetic species; it would not have temperature-dependent features in the MCD spectra. The spectra for both S=1/2 and S=3/2 $[4Fe\text{-}4S]^{+}$ cluster share positive features at approximately 380, 420, 520, and 740 nm. The main differentiator between S=1/2 and S=3/2 $[4Fe\text{-}4S]^{+}$ clusters is that S=1/2 has intense negative features between 600-700 nm. While there is significant noise in the higher energy region of the CotA MCD spectra, S=3/2 is the more likely spin state because of the magnetization saturation of the field dependent traces is due to the lack of negative features between 600-700 nm (**Fig. 6.9**).

The CotA MCD spectra are not reminiscent of other metalloproteins. The MCD spectra of CotA were compared to spectra of other chromophores to further validate the nomination of the Rieske [2Fe-2S] and [4Fe-4S] cluster as the likely CotA metal centers. There was not a spectral match that would indicate the MCD features arise from a chromophore other than a FeS cluster as summarized in **Table 6.3**.

Table 6.3. Discrepancies in MCD spectra for other metals and cofactors.

Chromophore	**Discrepancy**
Flavin	This is an organic cofactor; MCD spectra of flavin do not have temperature features.[23a]
F430	MCD spectra of F430 only have high energy temperature-dependent features (300-450 nm).[31]
Heme	Abs and MCD spectra of heme would have an intense Soret peak at approximately 400 nm.[32]
Cobalamin	Diagnostic Abs spectra with intense $\pi \rightarrow \pi^*$ feature; cobalamin species have unique MCD features that are absent in the spectra of CotA.[33]
Mononuclear-Mo	Mo(V) has weak d-d transitions in the MCD spectra; no shared MCD features with CotA.[34]
Mononuclear-Mn	Mn(III)/(II)-SOD have intense features at approximately 600 nm. Sign of MCD features do not match CotA.[35]
Mononuclear-Fe	No intense absorption spectra features; MCD spectra have multiple intense features in nIR.[36]
Mononuclear-Co	Abs spectra are different; MCD spectra contain an intense negative feature at approximately 600 nm. [37]
Dinuclear-Co	Convoluted Abs spectrum and MCD spectra has intense negative features 470-600 nm.[38]
Mononuclear-Ni	Ni-SOD (Ni(III)/(II)); MCD spectra has intense negative feature at 370 nm.[39]
Mononuclear-Cu	Copper centers have different absorption spectra; Cu metalloproteins have intense nIR features in MCD spectra.[14]
Mononuclear-Zn	As a d^{10} metal, the MCD spectra of Zn(II) would not have features in the visible region. [23a]

For example, a minor heme impurity could be the cause of this negative trough at 418 nm; however, this is unlikely as the Soret band is typically less broad than this CotA feature. Thus, the spectral interference from heme is unlikely even though the quartz discs used to build the home-made MCD cell have been used for MhuD, IsdG, and IsdI (non-canonical heme oxygenase enzymes studied by the Liptak group).

SECTION 6.4. CONCLUSIONS

The optical spectroscopy data we have collected strongly supports that CotA is a metalloprotein. This data does support the existence of an iron-sulfur cluster as opposed to another biologically relevant metal center or cofactor such as flavin. Based on the cataloged library of FeS cluster MCD spectra, the ISCs of CotA are most likely a Rieske [2Fe-2S] as well as a S=3/2 $[4Fe\text{-}4S]^{+}$. However, there are no published spectra that match that of CotA and the metal center cannot be definitively assigned. This could be due to a few reasons: (1) the sample is composed of monomeric, dimeric, trimeric, and possibly tetrameric CotA and their multiple clusters, (2) oxidation stress may have led to partial or complete degradation of the air-sensitive iron-sulfur clusters, or (3) a MCD spectra of the configuration of cluster contained by CotA has not been published.

The propensity for CotA to form dimers as opposed to higher order multimers may have biological significance. The dimers are quite stable as indicated by SDS-PAGE; as such, the dimers are likely coordinated through covalent bond(s). The formation of the dimer may bring the FeS cluster centers of each monomer close to each other and allow for two Rieske [2Fe-2S] clusters to form a [4Fe-4S] cluster. This has been reported previously where under reducing conditions, [2Fe-2S] clusters can combine.[40] The most common type of cluster interconversion is from a [2Fe-2S] to a [4Fe-4S] cluster.[15] This process has been observed in other iron-sulfur cluster proteins; however, dimerization of Rieske proteins has not been reported.

SECTION 6.5. FUTURE WORK

Variable temperature-variable field (VTVH) MCD is an accessible technique for the Liptak group that could be used to assign the spin state. Magnetization saturation curves that nest (lay on top of each other) would indicate a spin state of S>1/2 while saturation curves that completely overlay would determine that the cluster has a spin state of S=1/2. This negative feature was more challenging to assign, as intense negative features were found at different energies for these ISC; 418 nm is slightly higher in energy as compared to the published value of this Rieske feature. However, we do not believe that the magnetization saturation curves would be derived from a single ISC and be challenging to derive information pertaining to spin state of the major cluster. EPR spectroscopy would provide the spin-state information of a paramagnetic ISC.

Unfortunately, our collaborators in the Shen lab were unable to reproduce the phenotype of the -*cotA* knockout from the Cutting group.[7, 12] One proposed reason is that the different strains of *C. difficile* used could have an impact on whether CotA is an integral part of spore formation. To determine this, many strains of *cotA*$^{-}$ would have to be grown by our collaborators. The spore morphology of each mutant, as well as their sensitivity to lysozyme and ethanol, would have to be investigated for each strain of the bacterium. This is an intensive undertaking. The likely conclusion that would be reached from that experiment is that CotA is only involved in the spore formation for some strains of *C. difficile.* A therapy that targets only certain strains of *C. difficile* would not be effective for all infected patients nor would it be a marketable therapy.

These data may indicate that CotA may not be as important in the formation of the endospore in *C. difficile* as the Cutting groups data had implied. The malformed spore coat observed in the Cutting group may have been specific to the strain of *C. difficile* they used to conduct their investigations (CD630). The Cutting group noted that *C. difficile* spores grown in the laboratory environment have a morphology that differs from native *C. difficile*.[7] Based on this data, we no longer believe targeting the *cotA* gene would be an effective course of action to combat the increasing presence of antibiotic resistant strains of *C. difficile*.

SECTION 6.6. REFERENCES

1. Kelly, C. P.; LaMont, J. T., *Annu. Rev. Med.* **1998,** *49*, 375-390.

2. Henriques, A. O.; Moran, C. P., Jr., *Annu. Rev. Microbiol.* **2007,** *61*, 555-588.

3. Borriello, S. P., *Ann. Med.* **1990,** *22*, 61-67.

4. Riegler, M.; Sedivy, R.; Pothoulakis, C.; Hamilton, G.; Zacherl, J.; Bischof, G.; Cosentini, E.; Feil, W.; Schiessel, R.; LaMont, J. T., *J. Clin. Invest.* **1995,** *95*.

5. Paredes-Sabja, D.; Shen, A.; Sorg, J. A., *Trends Microbiol.* **2014**, 206-416.

6. Rubin, M. S.; Bodenstein, L. E.; Kent, K. C., *Dis. Colon. Rectum.* **1995,** *38*.

7. Permpoonpattana, P.; Phetcharaburanin, J.; Mikelsone, A.; Dembek, M.; Tan, S.; Brisson, M. C.; La Ragione, R.; Brisson, A. R.; Fairweather, N.; Hong, H. A.; Cutting, S. M., *J. Bacteriol.* **2013,** *195*, 1492-1503.

8. Viscidi, R.; Willey, V.; Bartlett, J. G., *Gastroenterology* **1981,** *81*, 5-9.

9. Baines, S. D.; O'Connor, R.; Saxton, K.; Freeman, J.; Wilcox, M. H., *J. Antimicrob. Chemother.* **2009,** *63*.

10. Gerding, D. N.; Muto, C. A.; Owens, R. C., Jr., *Clin. Infect. Dis.* **2008,** *46 Suppl 1*, S43-49.

11. Deakin, L. J.; Clare, S.; Fagan, R. P.; Dawson, L. F.; Pickard, D. J.; West, M. R.; Wren, B. W.; Fairweather, N. F.; Dougan, G.; Lawley, T. D., *Infect. Immun.* **2012,** *80*, 2704-2711.

12. Permpoonpattana, P.; Tolls, E. H.; Nadem, R.; Tan, S.; Brisson, A.; Cutting, S. M., *J. Bacteriol.* **2011,** *193*, 6461-6470.

13. Passerini, A.; Lippi, M.; Frasconi, P., *Nucleic Acids Res.* **2011,** *39*, W288-292.

14. Liu, J.; Chakraborty, S.; Hosseinzadeh, P.; Yu, Y.; Tian, S.; Petrik, I.; Bhagi, A.; Lu, Y., *Chem. Rev.* **2014,** *114*, 4366-4469.

15. Fee, J. A.; Findling, K. L.; Yoshida, T.; Hille, R.; Tarr, G. E.; Hearshen, D. O.; Dunham, W. R.; Day, E. P.; Kent, T. A.; Munck, E., *J. Biol. Chem.* **1984,** *259*, 124-133.

16. Hanzelmann, P.; Hernandez, H. L.; Menzel, C.; Garcia-Serres, R.; Huynh, B. H.; Johnson, M. K.; Mendel, R. R.; Schindelin, H., *J. Biol. Chem.* **2004,** *279*, 34721-34732.

17. Shen, A.; Lupardus, P. J.; Morell, M.; Ponder, E. L.; Sadaghiani, A. M.; Garcia, K. C.; Bogyo, M., *PLoS One* **2009,** *4*, e8119.

18. Studier, F. W., *Protein Expr. Purif.* **2005,** *41*, 207-234.

19. Lockhart, C. L.; Conger, M. A.; Pittman, D. S.; Liptak, M. D., *J. Biol. Inorg. Chem.* **2015,** *20*, 757-770.

20. Crack, J. C.; den Hengst, C. D.; Jakimowicz, P.; Subramanian, S.; Johnson, M. K.; Buttner, M. J.; Thomson, A. J.; Le Brun, N. E., *Biochem.* **2009,** *48*, 12252-12264.

21. Ollagnier-de-Choudens, S.; Mattioli, T.; Takahashi, Y.; Fontecave, M., *J. Biol. Chem.* **2001,** *276*, 22604-22607.

22. Rouault, T. A.; Tong, W. H., *Nat. Rev. Mol. Cell Biol.* **2005,** *6*, 345-351.

23. (a) Que, L., *Physical Methods in Bioinorganic Chemistry: Spectroscopy and Magnetism*. University Science Books: Sausalito, Calif., 2000; (b) Collins, D. J.; Zhou, H.-C., Iron-sulfur models of protein active sites. In *Encyclopedia of Inorganic Chemistry*, 2nd ed.; Wiley: 2005.

24. (a) Stephens, P. J.; Thomson, A. J.; Dunn, J. B. R.; Keiderling, T. A.; Rawlings, J.; Rao, K. K.; Hall, D. O., *Biochem.* **1978,** *17*, 4770-4778; (b) Stephens, P. J.; Thomson, A. J.; Keiderling, T. A.; Rawlings, J.; Rao, K. K.; Hall, D. O., *Proc. Natl. Acad. Sci. USA* **1978,** *75*, 5273-5275.

25. Pandelia, M. E.; Lanz, N. D.; Booker, S. J.; Krebs, C., *Biochim. Biophys. Acta.* **2015,** *1853*, 1395-1405.

26. (a) Johnson, C. E.; Bray, R. C.; Cammack, R.; Hall, D. O., *Proc Natl Acad Sci USA* **1969,** *63*, 1234-1238; (b) Rouault, T. A., *Nat. Chem. Biol.* **2015,** *11*, 442-445.

27. Hanson, G.; Berliner, L.; SpringerLink, Metals in Biology Applications of High-Resolution EPR to Metalloenzymes. In *Biological Magnetic Resonance*, [Online] Springer New York,: New York, NY, 2010; pp. XIX, 419p.

28. Subramanian, S.; Duin, E. C.; Fawcett, S. E.; Armstrong, F. A.; Meyer, J.; Johnson, M. K., *J. Am. Chem. Soc.* **2015,** *137*, 4567-4580.

29. Duderstadt, R. E.; Brereton, P. S.; Adams, M. W.; Johnson, M. K., *FEBS Lett.* **1999,** *454*, 21-26.

30. Mapolelo, D. T.; Zhang, B.; Naik, S. G.; Huynh, B. H.; Johnson, M. K., *Biochem.* **2012,** *51*, 8056-8070.

31. Hamilton, C. L.; Scott, R. A.; Johnson, M. K., *J. Biol. Chem.* **1989,** *264*, 11605-11613.

32. Beinert, H.; Holm, R. H.; Munck, E., *Science* **1997,** *277*, 653-659.

33. Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C., *J. Am. Chem. Soc.* **2003,** *125*, 5897-5914.

34. Hernandez-Marin, E.; Seth, M.; Ziegler, T., *Inorg. Chem.* **2009,** *48*, 2880-2886.

35. Jackson, T. A.; Gutman, C. T.; Maliekal, J.; Miller, A. F.; Brunold, T. C., *Inorg. Chem.* **2013,** *52*, 3356-3367.

36. Solomon, E. I.; Zhang, Y., *Acc. Chem. Res.* **1992,** *25*, 343-352.

37. Larrabee, J. A.; Alessi, C. M.; Asiedu, E. T.; Cook, J. O.; Hoerning, K. R.; Klingler, L. J.; Okin, G. S.; Santee, S. G.; Volkert, T. L., *J. Am. Chem. Soc.* **1997,** *119*, 4182-4196.

38. Johansson, F. B.; Bond, A. D.; Nielsen, U. G.; Moubaraki, B.; Murray, K. S.; Berry, K. J.; Larrabee, J. A.; McKenzie, C. J., *Inorg. Chem.* **2008,** *47*, 5079-5092.

39. Fiedler, A. T.; Bryngelson, P. A.; Maroney, M. J.; Brunold, T. C., *J. Am. Chem. Soc.* **2005,** *127*, 5449-5462.

40. Duin, E. C.; Lafferty, M. E.; Crouse, B. R.; Allen, R. M.; Sanyal, I.; Flint, D. H.; Johnson, M. K., *Biochem.* **1997,** *36*, 11811-11820.

CHAPTER 7: CBA1: A UNIQUE COBALAMIN ACQUISITION PROTEIN FROM *PHAEODACTYLUM TRICORNUTUM*

SECTION 7.1 INTRODUCTION

The diatom, *Phaeodactylum tricornutum*, encodes for a putative cobalamin acquisition enzyme. *P. tricornutum* is an aquatic Arctic diatom involved in the global carbon recycling. The genome of this microalgae was sequenced in 2007; *P. tricornutum* expresses both forms of methionine synthase.[1] The proposed function of MMCM in diatomic algae may be involvement in the citric acid cycle or fatty acid metabolism, *P. tricornutum* and other diatoms encode the AdoCbl-dependent enzyme.[2,3] While the diatom does not have an absolute cobalamin requirement, cobalamin-replete environments adversely affect algal growth.[2] The micronutrient is found in dilute concentrations in the ambient aquatic environment of *P. tricornutum*.[2]

When starved of cobalamin in a laboratory environment, the Saito group (Woods Hole Oceanography Institute, Woods Hole, MA) observed a marked increase in expression of a single protein: CBA1.[2] Low cobalamin did not affect MMMC expression.[2] To better understand the role of CBA1 in cobalamin uptake, a series of transgenic *P. tricornutum* cell lines overexpressing CBA1 (CBA1-OE1 and CBA1-OE2) were monitored for their cobalamin uptake and compared to wild-type *P. tricornutum* and a cell line overexpressing urease (Urease-OE1) (**Fig. 7.1**). The correlation between the overexpression of CBA1 and cobalamin uptake confirmed CBA1 was involved cobalamin uptake in *P. tricornutum*.

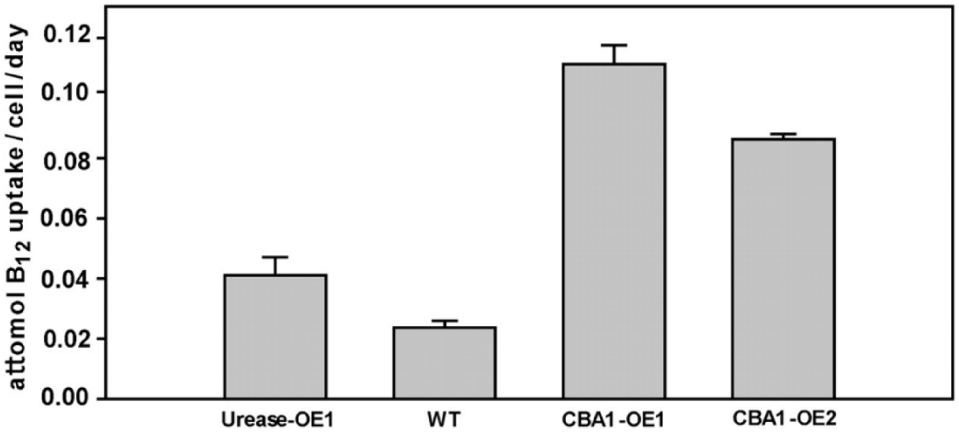


Figure 7.1. Both wild-type (WT) and Urease-OE1 (protein unrelated to cobalamin uptake) were used as controls. The algae were stored in a cobalamin-replete media and cobalamin uptake was monitored over 24 hours. Figure taken from reference 9.

The CBA1 protein was engineered to have a yellow fluorescent protein tag *in vivo* understand how CBA1 acquired cobalamin from the aquatic environment of *P. tricornutum*. The N-terminal signal peptide was found in the sequence of CBA1 that would direct the diatom to excrete CBA1.[2] CBA1 was found localized on the outside of the diatom (**Fig. 7.2**).

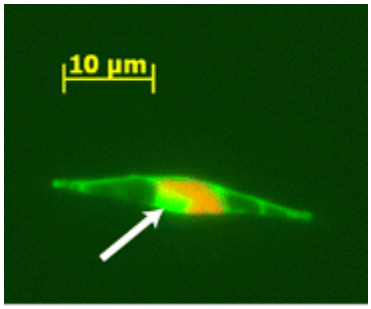


Figure 7.2. Epifluorescent microscopy image of *P. tricornutum* showing CBA1 congregating on the outside of the diatom. CBA1 is fused to yellow fluorescent protein and emits green light upon excitation. The white arrow is highlighting a second site of CBA1 aggregation in an organelle next to the chloroplast. Figure taken from reference 9.

The putative role of CBA1 is to capture cobalamin from the seawater and bring the micronutrient into the cell.[2] A response system comprised of several proteins, including CBA1, allows *P. tricornutum* to thrive in cobalamin-replete seawater. To counteract the deleterious effects of the inefficient cobalamin-independent methionine synthase as a response to cobalamin scarcity, the production of these proteins is enhanced in the algae: CBA1, MetE (cobalamin-independent methionine synthase), MetK (S-adenosylmethionine synthase) and ThiC (thiamine biosynthesis protein).[2] Together these four enzymes lessen the burden of cobalamin scarcity: increasing cobalamin acquisition sites (CBA1), lowering the need of cobalamin (MetE), and mitigating the effects caused by the decrease in methionine synthase activity with the less efficient MetE (ThiC and MetK). This response system is advantageous for the diatom as it increases the likelihood of acquiring cobalamin from the environment.

The mechanism of cobalamin acquisition by CBA1 in *P. tricornutum* has yet to be determined. Better understanding of cobalamin binding by CBA1 will offer insight into cobalamin acquisition in aquatic diatoms. Ultimately, this will provide information to better understand how cobalamin scarcity affects marine phytoplankton and how this influences their integral role in global carbon cycling.

SECTION 7.2. METHODS

Recombinant expression of CBA1 in *Escherichia coli*. Recombinant DNA was designed in collaboration with Mak Saito from Woods Hole Oceanography Institute. Recombinant CBA1 was cloned into the pDEST17 vector then transformed into a BL21-GOLD (DE3) *E. coli* cell line for expression. The gene was designed to include a N-terminal His_6-tag and a tobacco etch virus protease (TEV) cleavage site to optimize and simplify the purification procedure (**Fig. 7.3**).

XXX	HHHHHH	$(XXX)_n$	ENLYFQG	XXX	DEPPACLTST…
Linker	His_6-tag	Linker	TEV cleavage site	Linker	CBA1

Figure 7.3. CBA1 gene structure in the pDEST17 vector.

Upon the transformation of the plasmid into the *E. coli* cell line, glycerol stocks were flash frozen in liquid nitrogen and stored at -80 °C. The transformation was confirmed by sequencing extracted DNA at the Vermont Cancer Center (**Fig. F.1**). With the His_6-tag and TEV cleavage site, the molecular weight of CBA1 is 51 kDa. After the TEV cleavage reaction, the molecular weight of cleaved CBA1 is 47 kDa. TEV cleaves the protein between the glutamine (Q) and glycine (G) residues in the cleavage site.

Growth media and parameters for CBA1 encoded *E. coli*. A lysogeny broth (LB)-agar plate was brought to room temperature and a 50 µL aliquot of 100 mg/mL ampicillin was evenly distributed and left to soak into the agar for 30 min. A sterile inoculating loop was used to spread the frozen *E. coli* BL21-GOLD (DE3) cells onto the LB agar plate. The plate was aerobically incubated overnight at 37 °C. The following

morning, a single colony was transferred from the LB-agar plate to 20 mL of sterile LB media with a 20 μL aliquot of 100 mg/mL ampicillin. This mini-growth was shaken at 225 rpm at 37 °C overnight in an aerobic environment.

If the mini-growth was cloudy the next morning, an aliquot would be transferred to 1 L of media in a Fernbach flask. Two different recipes for growth media were compared to optimize the expression of CBA1: LB and terrific broth (TB). For every 1 L of LB media, 10 g tryptone, 10 g NaCl, and 5 g yeast were added to 1 L of MilliQ water. This media was autoclaved for sterilization. A 10 mL aliquot of cloudy mini-growth was sterile filtered into 1 L of LB media along with 1 mL of 100mg/mL ampicillin. For every 1 L of TB, 24 g yeast extract, 12 g tryptone, and 4 mL glycerol were added to 880 mL of MilliQ water and autoclaved. A 20 mL aliquot of cloudy mini-growth, 1.5 mL 100 mg/mL ampicillin, and 10 mL potassium phosphate buffer at pH 7.4 (0.17 M monobasic potassium phosphate, 0.74 M dibasic potassium phosphate solution autoclaved for sterility) were sterile filtered into the room temperature TB.

Expression of the cobalamin acquisition protein in an *E. coli* cell line. The promotor sequences in the pDEST17 vector containing the CBA1 gene was compatible with three induction methods. Induction by isopropyl β-D-1-thiogalactopyranoside (IPTG), induction by L-arabinose, and auto-induction were tested. IPTG and auto-induction target the lacI promoter while L-arabinose targets the P_{BAD} promoter.[4,5,6] To optimize induction protocol for CBA1, the methods were compared through sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). A molecular weight

ladder was used to identify the band corresponding to His_6-tagged CBA1 at 51 kDa (**Fig. F.2**).

SDS-PAGE was used to identify the proteins by their molecular weights in each wash. SDS-PAGE denatures the proteins and the lightest proteins travel farthest down the gel. The polyacrylamide gel was stained with GelCode Blue (ThermoFisher). The molecular weight of the protein bands was determined by a molecular weight ladder (**Fig. F.2**).

The methodology for IPTG induction was adapted from a previously published procedures.[4] Ten mL aliquot of the mini-growth was sterile-filtered into 1 L of sterile LB media (referred to as a "large growth" in the chapter) which was then grown at 37 °C and shaken at 225 rpm. The optical density at 600 nm (OD_{600}) was monitored by UV-vis absorption (Abs) spectroscopy. Once an OD_{600} between 0.5 and 0.8 was reached, 10mL of 100mM IPTG was sterile filtered into the large growth and growths were incubated for 3 hours under the aforementioned conditions. Following the incubation period, the cells were harvested by centrifugation at 10,000 x*g* at 4 °C for 15 min. The cell pellet was removed and stored at -80 °C in 50 mL conical tubes.

The methodology of L-arabinose induction was adapted from a previously published procedures.[5] A ten mL aliquot of the cloudy mini-growth, grown as previously described, was sterile filtered into 1 L of LB media with 1 mL of 100mg/mL ampicillin. The culture was grown at 37 °C while shaking at 225 rpm until the OD_{600} reached 0.4 a.u. then 10 mL of 20 % (*w/w*) L-arabinose was sterile filtered into the large growth. The culture was grown for 3 hours while shaking at 225 rpm at 37 °C. After the period of induced

growth, the culture was divided into equal aliquots for centrifugation at 15,000 x*g* and 4 °C for 30 min. The harvested pellet was stored in 50 mL conical tubes at -80 °C until required for CBA1 purification.

The methodology of auto-induction was adapted from a previously published procedure.[6] The min-growth was grown as previously described. Aliquots of 20 mL eacg of mini-growth and sugar mix (25 % glycerol, 2.5 % glucose, 10 % α-lactose monohydrate solution) were sterile filtered into 1 L of room temperature TB. The inoculated media was incubated at 37 °C while shaking at 225 rpm for 16 hours. The media was then divided into equal aliquots for centrifugation at 10,000 x*g* at 4 °C for 15 min. The cell pellets were kept at -80 °C in 50 mL conical tubes for long-term storage.

Lysis of *E. coli* cells to release over-expressed CBA1. Two methods of lysis were compared to extract the maximum amount of CBA1 from *E. coli*. Lysis destroys the cell membrane and releases the protein of interest along with natively expressed proteins. Two common methods of lysis were compared to optimize protein yield by effectively lysing the cell pellets: enzymatic lysis and sonication.

To enzymatically lyse the *E. coli* cells, a cell pellet was thawed and resuspended with lysis buffer (see **Table F.1**). To the resuspended pellet, 25 mg chicken egg white lysozyme, 5 μL DNase I, and 250 μL 0.01 M PMSF were added, the solution was incubated for one hour, then centrifuged at 15,000 x*g* at 4 °C for 30 min. The lysate was separated from the cell pellet and insoluble biomaterials were removed by vacuum filtration.

To lyse the *E. coli* cells by sonication, the pelleted cells were brought to room temperature in water then resuspended in lysis buffer (see **Table F.1**). The cells were lysed

on ice at 70 % sonication duty for four cycles of one minute of sonication followed by a one minute rest period. A tapped step horn with threaded body sonifier with a 13 mm tip was used to sonicate CBA1-containing lysate. The lysate was kept on ice during sonication to prevent degradation. To separate the insoluble debris, the lysate was centrifuged at 15,00 x*g* at 4 °C for 30 min. The supernatant was filtered by vacuum filtration before it was used for protein purification. Some variations of sonication lysate were preceded by a short (30 min) lysozyme incubation period to enhance lysis.

Purification of CBA1. Affinity chromatography was used to isolate His_6-tagged CBA1. Ni(II) nitrilotriacetic acid (NTA) affinity columns have a higher affinity, thus a higher binding capacity, for all proteins. The Co(II)-NTA has a high specificity for polyhistidine tags; while less protein is bound to the column, it is more likely to be the His_6-tagged protein of interest.

Herein is discussed the general procedure for purification of CBA1. The NTA affinity column was equilibrated with equilibration buffer (**Table F.2**) before the CBA1-containing lysate was loaded onto the column. The lysate-loaded column was washed with a series of buffers with increasing concentrations of imidazole ranging from 0 mM to 400 mM (**Table F.3**). The column was then washed with regeneration buffer (20 mM sodium acetate, 100 mM NaCl at pH 5.0) to remove any tightly bound proteins.

In attempt to increase the yield of the protein of interest, ammonium sulfate precipitation method was adapted from published reports.[7] A step-wise increase of ammonium sulfate was used to determine the ideal concentration to pull CBA1 out of solution while native proteins remained in solution in a process called “salting-out”. The

amount of ammonium sulfate was increased from 1.0 M to 4.0 M in 0.5 M increments to determine the optimal conditions for salting-out CBA1 (**Fig. F.7**). The lysate with ammonium sulfate was incubated for 30 min at room temperature then centrifuged and analyzed for cell pellet formation. Pellets were resuspended in 50 mM Tris buffer at pH 7.4. The concentration of protein was determined by a NanoDrop Spectrophotometer (**Table F.4**).

SECTION 7.3. RESULTS

Optimal induction method for over-expression of CBA1 in *E. coli*. SDS-PAGE gels were compared to identify the best expressed procedure. IPTG induction did not lead to the overexpression of CBA1 as compared to native proteins (**Fig. F.3**). CBA1 may have been overexpressed with the autoinduction method but the expression of native proteins was also enhanced (**Fig. F.4**). While a dark band at the appropriate molecular weight for CBA1 was found in the lysate of autoinduced media, CBA1 was not present in any elution wash (**Fig. F.4**). A dark band at approximately 51 kDa was detected in the L-arabinose induced lysate; this was more apparent in the 50 mM imidazole wash (lane 2) in the SDS-PAGE gel (**Fig. F.5**). Induction by L-arabinose led to selective overexpression of CBA1 and the protein was present in elution washed.

Ideal lysis method for over-expression of CBA1 in *E. coli*. Lysate from enzymatic lysis and sonication were compared by SDS-PAGE. Compared to enzymatic lysis, sonication was far more effective at breaking open the *E. coli* cell wall and releasing the protein of interest (**Fig. F.4 and F.6**). Sonication was used in all purification attempts.

Purification procedure to optimize the yield of CBA1 expressed in *E. coli*. Once CBA1 expression was sufficient with L-arabinose induction, several permutations of the general purification procedure were used to isolate pure CBA1 from the lysate. The most successful lysis attempt used a pre-sonication incubation with lysozyme followed by sonication (**Figs. F.8** and **F.9**). However, three challenges plagued purification that prevented sufficient CBA1 yield and purity: low expression, poor solubility and improper folding.

CBA1 was not found at high concentrations in the cell lysate; this may be indicative of poor solubility or expression. Three expression methodologies were compared and L-arabinose was shown to lead to the highest expression of CBA1 (**Fig. F.5**). Detergent (Triton X-100) was added in an attempt to increase the solubility of CBA1 in solution. The addition of the nonionic detergent did not significantly increase the solubility of CBA1. Salting-out with ammonium sulfate was attempted take advantage of the low solubility of CBA1. A high ammonium sulfate concentration (3.5 M) led to the selective precipitation of CBA1 while native proteins remained in solution. This pellet of CBA1 was resuspended in equilibration buffer but the protein was still relatively insoluble in the buffer conditions (**Fig. F.9**).

The presence of CBA1 in 50 mM elution washed indicated that the protein was not binding tightly to the Ni(II)-NTA resin (**Figs. F.5- F.9**). This could be due to the improper folding of the protein or the His_6-tag was incorporated into the protein fold. To maintain proper folding of CBA1 during lysis, glycerol was added to the solution. Glycerol did not greatly increase the yield of CBA1 for spectroscopic studies.

SECTION 7.4. DISCUSSION

Purification attempts of CBA1 did not lead to enough isolated protein for spectroscopic investigation. The fractions containing CBA1 were not pure enough for cobalamin binding studies. Putative toxicity and low solubility were problematic for CBA1 expression. Poor expression is the most likely cause for the low yield of CBA1.

The expression of CBA1 in the bacterial host, *E. coli*, may have been unsuccessful because *P. tricornutum* is an algal organism. As a bacterium, *E. coli* may not be equipped with the cellular machinery found in Antarctic algae to properly translate or fold CBA1. *E. coli* is grown ideally at 37 °C which may not be conducive for the Antarctic protein. An expression host for algal proteins or membrane proteins may be better equipped for CBA1 expression but the solubility problems of CBA1 would remain.

Low expression may be because CBA1 could be toxic to the host.[8] If CBA1 is removing cobalamin from the cellular environment, *E. coli* would be able to survive because it encodes for both MetE and MetH.[9] However, MetE is less efficient and solely relying on the cobalamin-independent methionine synthase enzyme would adversely affect growth. For insoluble proteins, a fusion protein tag may increase solubility to ease extraction of the protein of interest from the lysate.

There are different vectors, such as pET, available that could had been used for cloning CBA1 instead of the pDEST17 vector. The vector could also be transformed into a cell line other than BL21-GOLD (DE3). Expression enhancement would be unlikely following the transformation of the CBA1 gene into a different vector or *E. coli* cell line.

Poor protein yields could be due to low solubility of CBA1. The N-terminal signal peptide in the CBA1 sequence directed the algae to excrete the protein; CBA1 could be associated with the outside of the *E. coli* cell membrane.[2] If CBA1 was insoluble in the lysis buffer, it would be discarded after centrifugation along with the cell membrane pellet. Large protein adducts can be added to proteins of interest to increase solubility.[10] However, any protein tag would have to be cleaved prior to spectroscopic investigation of cobalamin binding. Incomplete or incorrect folding of CBA1 in *E. coli* could be the cause for the low solubility. The most successful protocol for purifying other proteins from *P. tricornutum* by our collaborators required a high concentration of urea (can increase the solubility of the protein). Ultimately, both poor expression and solubility of CBA1 prevented spectroscopic investigation of cobalamin binding.

SECTION 7.5 CONCLUSIONS AND FUTURE WORK

Challenges with purifying CBA1 lies in poor expression as well as solubility. Currently, attempts to improve purification such that CBA1 was of high yield and purified without impurities have been halted. Discussions with collaborators led to the determined that an optimized purification protocol could not produce enough purified protein to allow cobalamin binding by CBA1 to be studied spectroscopically.

SECTION 7.6. CHAPTER 7 REFERENCES

1. De Martino, A.; Meichenin, A.; Shi, J.; Pan, K.; Bowler, C., *Journal of Phycology* **2007,** *43*, 992–1009.

2. Bertrand, E. M.; Allen, A. E.; Dupont, C. L.; Norden-Krichmar, T. M.; Bai, J.; Valas, R. E.; Saito, M. A., *Proc. Natl. Acad. Sci. USA* **2012,** *109*, E1762-1771.

3. Croft, M. T.; Warren, M. J.; Smith, A. G., *Eukaryotic Cell* **2006,** *5*, 1175-1183.

4. Sivashanmugam, A.; Murray, V.; Cui, C.; Zhang, Y.; Wang, J.; Li, Q., *Protein Sci.* **2009,** *18*, 936-948.

5. Khlebnikov, A.; Risa, O.; Skaug, T.; Carrier, T. A.; Keasling, J. D., *J. Bacteriol.* **2000,** *182*, 7029-7034.

6. Studier, F. W., *Protein Expr. Purif.* **2005,** *41*, 207-234.

7. Wingfield, P., *Curr. Protoc. Protein Sci.* **2001,** *Appendix 3*, Appendix 3F.

8. Rosano, G. L.; Ceccarelli, E. A., *Front. Microbiol.* **2014,** *5*, 172.

9. Pejchal, R.; Ludwig, M. L., *PLoS Biol.* **2005,** *3*, e31.

10. Shen, A.; Lupardus, P. J.; Morell, M.; Ponder, E. L.; Sadaghiani, A. M.; Garcia, K. C.; Bogyo, M., *PLoS One* **2009,** *4*, e8119.

CHAPTER 8. PROBING THE SUPERNUCLEOPHILICITY OF COBALT TETRAPYRROLES WITH COBALT-59 NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY.

SECTION 8.1. INTRODUCTION

Market for rapid spectral analysis of cobalt tetrapyrroles. The need for an extensive library of cobalt-59 nuclear magnetic resonance (NMR) spectra is justified by the vast scope of the literature concerning various cobalt tetrapyrroles and their applications.[1] Cobalt tetrapyrroles are ubiquitous cofactors in biological systems and have been shown to participate in many unique catalytic cycles of both biological and synthetic significance. The applications of cobalt tetrapyrroles include decontamination (as a dehalogenation catalyst) as well as a highly attractive alternative energy source (as a hydrogen evolution catalyst).[2,3] The catalytic prowess of the cobalt tetrapyrroles is rooted in the accessible Co(I) oxidation state; the super-reduced cobalamin, cob(I)alamin, has been described as having supernucleophilic character as is the most powerful biological nucleophile to date.[4] The supernucleophilicity of the cob(I)alamins and other tetrapyrroles is due to the out-of-plane occupied d_z^2 orbital.[5] The equatorial ligand influences the nucleophilicity by altering the electron density of the metal center.[5]

Cobalt-59 NMR spectroscopy provides insight into a variety of catalyses by cobalt tetrapyrroles. The ^{59}Co chemical shift of the cobalt tetrapyrroles provides information about the σ-bond interactions between the metal and the axial ligand. The d_z^2 orbital is the lowest unoccupied Co-based molecular orbital (LUMO) in Co(III) tetrapyrroles and the highest occupied Co-based molecular orbital (HOMO) in Co(I) tetrapyrroles. This orbital has a direct impact on the nucleophilicity of the cobalt center as well as the chemical shift. The source of the chemical shift is related to the ligand strength and type, which influence the d-orbital energies of the metal center.[1a] Therefore, ^{59}Co

solution state NMR will be a direct probe of the supernucleophilicity of cobalt tetrapyrrole catalysts.

For the diamagnetic low-spin Co(I) and Co(III) species, the chemical shift will be dictated by the paramagnetic contributions (σ_p).[6] Paramagnetic contributions are influenced by the shape of ground state orbital as well as the energy gap of the HOMO and LUMO of the probed nuclei.[6] The ground state orbital must non-spherical or it will not be able to rotate upon application of an external magnetic field. The magnetic dipole-allowed excited states will mix upon the introduction of the magnetic field. The rotation of orbitals increases the effective magnetic field which causes the nuclei to resonate at a higher frequency.[7]

The discussion of the cobalt-59 chemical shifts in this chapter will focus on the energy gap between the highest occupied and the lowest unoccupied Co d-orbitals. A small energy gap between the ground and excited states imposes less of an energetic penalty which manifests as more of a downfield shift. Therefore, the paramagnetic contribution is inversely related to the ligand field splitting (LFS) of the cobalt nuclei (**Eqn. 8.1**). For example, a Co(III) tetrapyrrole with a large ligand field splitting, the chemical shift will be found in relatively upfield from a small paramagnetic contribution. Increasing the binding affinity of the varied ligand (consequently increasing the LFS) results in an upfield chemical shift.[1a] When a single ligand is different between two cobalt tetrapyrroles, an upfield chemical shift should follow the following simplified trend based on the spectrochemical series of the ligating atom: (downfield) O < N < C (upfield).

$$\sigma_p \propto {}^{1}/_{LFS} \qquad \textbf{Equation 8.1}$$

Based on the ligands' binding affinities, CNCbl should have a smaller HOMO-LUMO gap compared to stronger alkyl ligands MeCbl and AdoCbl.[1a] Of these three biologically relevant cobalamins, CNCbl should have the largest paramagnetic contribution and subsequently the most downfield chemical shift. As ligand field splitting increases with oxidation state, a Co(III) tetrapyrrole will have a larger HOMO-LUMO gap and upfield shift compared to a Co(I) tetrapyrrole (**Fig 8.1**).

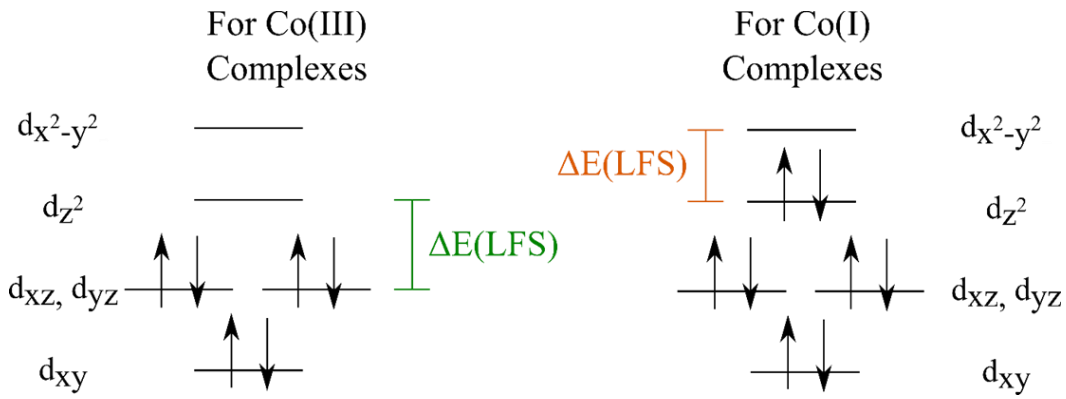


Figure 8.1. Ligand field splitting of low-spin Co(III) and Co(I) tetrapyrrole complexes.

Accessibility of Cobalt-59 NMR spectroscopy for variety of disciplines. Cobalt-59 NMR can be used as an extensive fingerprinting tool for many cobalt complexes including, but not limited to, cobalt tetrapyrroles. The natural abundance of this quadrupolar nuclei is 100% so isotopic-labelling is not required to enhance the signal.[8] The ^{59}Co nucleus resonance is similar to that of ^{13}C; the resonance frequency for cobalt-59 is 118.071 MHz and 125.721 MHz for carbon-13 at 11.744 T.[9] If the ^{13}C nuclei can be detected by a NMR spectrometer with a tunable broadband probe, it can be optimized for the ^{59}Co nuclei.

Challenges with cobalt-59 NMR data collection. Due to the very wide features and 20,000 ppm range for the resonance, it is difficult to detect cobalt nuclei. The rapid relaxation is due to the large nuclear spin, I=7/2, of ^{59}Co.[1a] Typically the relaxation time is less than one second for samples in solution; this quadrupolar relaxation is the primary cause for broad spectral peaks. Peak widths can be as wide as the entire sweep width for a typical ^{13}C spectra (approximately 250 ppm).[10] The line widths can be used to pull out information pertaining to the electronic field gradients.[1a] The electronic field gradient provides information about symmetry local to the cobalt center, short and long range structure, as well as their conformational heterogeneity.

SECTION 8.2. METHODS

Sample prep. Materials were purchased from Fisher unless otherwise noted. Cobalamin samples were studied at a concentration of 3-4 mM in D_2O (Cambridge Isotope Laboratories) or a biologically relevant buffer at physiological pH (10% D_2O). To build an extensive library of solution-state cobalt tetrapyrroles, spectra were collected in a commonly used biological buffer (Tris) at a physiological relevant pH. Tris buffer was chosen because of its low conductivity, 0.06 mS/cm, which will only minimally decrease the signal.[11]

Cyanocobalamin (CNCbl or vitamin B12) was studied as purchased. Aquocobalamin (H_2OCbl or Vitamin B12a) and hydroxocobalamin were also studied as purchased from Sigma-Aldrich (HOCbl·HCl or HOCbl·HOAc) and the pH was adjusted to protonate the β-axial ligand; the pKa of the water ligand of H_2OCbl is 7.8.[12] Imidazolylcobalmin (ImCbl) was prepared by adding imidazole to aquocobalamin in slight excess (1.3 imidazole: 1 H_2OCbl) as per the published protocol.[13] Bisimidazolylcobalmin (bisImCbl) was prepared by adding excess imidazole to aquocobalamin as per the published protocol.[14] Dicyanocobalamin (diCNCbl) was prepared by adding excess potassium cyanide to CNCbl in 50 mM Tris buffer at pH 8.0.[15] Cyanatocobalamin (OCNCbl) was prepared by adding excess potassium cyanate to H_2OCbl in 50 mM Tris buffer at pH 8.0.[16] Cob(I)alamin was prepared anaerobically by incubating aquocobalamin in 50 mM Tris with granular zinc and ammonium chloride under a N_2 atmosphere.[17] Prepared samples were kept at 4 °C to prevent degradation; samples were warmed to room temperature prior to data collection.

UV-vis absorption (Abs) spectroscopy. The room-temperature Abs spectra of these cobalamins were collected with a Cary 100 Bio spectrophotometer set to scan from 800 to 300 nm with a scan rate of 300 nm/min and 0.5 nm data interval. Samples were diluted 100-fold to a final concentration of approximately 0.03 mM.

^{1}H NMR spectroscopy. 1D ^{1}H NMR were recorded in D_2O or buffer at physiological pH with 10% D_2O on a Varian Unity Inova 500 MHz NMR spectrometer equipped with an inverse triple-resonance probe. The standard ^{1}H NMR experiments (zg pulse program) utilized a 1.5 sec relaxation delay and a 1 sec acquisition time. The ^{1}H water saturation experiments (PRESAT) utilized a 50 msec τ-delay and a 50 msec acquisition time. All ^{1}H NMR and ^{1}H PRESAT data were referenced to internal sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS; Cambridge Isotope Laboratories) or to DSS via the residual water peak. The data were processed using the ACD/Labs NMR Processor with zero filling to 65536 points and no line broadening was applied to the data.

^{59}Co NMR spectroscopy pulse program. Solution-state ^{59}Co NMR spectra were collected on a Bruker Avance III 500 MHz NMR spectrometer equipped with a broadband detection probe. The antiring (aring) pulse program was used to detect ^{59}Co nuclei (**Fig. 8.2**). ^{59}Co NMR experiments utilized relaxation delay of 8 μsec, 1 msec dwell time, and 1 msec acquisition time. Between pulses, a 4 μsec delay was incorporated as a post-pulse delay (d13). 50,000 scans were collected in approximately 25 min. All ^{59}Co NMR were Ξ-referenced to the ^{1}H resonance of water (see **Chapter G** for more information).[18,19] The spectra were processed using ACD/Labs NMR Processor with 5000 Hz line broadening and zero filling to 65536 points.

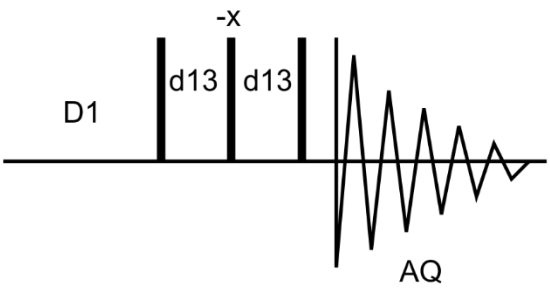


Figure 8.2. The aring pulse program used for ^{59}Co NMR data collection. D1 stands for pre-scan delay, d13 signifies post-pulse delay, AQ stands for acquisition time, and the thick bars (-x) stand for three -90° pulses. The nuclei then relax to the +z-plane during AQ.

Density functional theory (DFT) Calculations. The cobalamins were modeled by DFT calculations. The computations were performed with the ORCA 3.0.0 software package on the 380 node IBM Bluemoon cluster at the Vermont Advanced Computing Core. The restricted Kohn−Sham calculations utilized a Perdew−Burke−Ernzerhof (PBE) generalized gradient functional with a triple-zeta basis set with tight SCF convergence criteria.[20,21] The PBE functional has been successfully used in the past for simplified cobalamin models.[22] All computations used simplified cobalamins as model structures prepared in ArgusLab (Planaria Software) from reported crystal structures of CNCbl (CCDC: 766943). The simplified cobalamin truncated all alkyl chains to methyl groups and removed the nucleotide loop anchoring 5,6-dimethylbenzimidazole (DMB) to the corrin ring (**Fig. 8.3**). This simplified model of the tetrapyrrole is reminiscent of models studied previously and is a compromise between computational costs and cofactor complexity; prior models replace all substituents on the corrin ring with hydrogen atoms.[23] Löwdin orbitals were plotted to assign the Co d-orbital contribution to each MO. The utility

program orca_plot was used to generate molecular orbitals for each tetrapyrrole. An isodensity of ±0.03 a.u. was used to plot these orbitals using gOpenMol. The molecular orbitals plotted were the five MOs lower in energy than the HOMO, the HOMO, the LUMO, and the five MOs higher in energy than the LUMO.

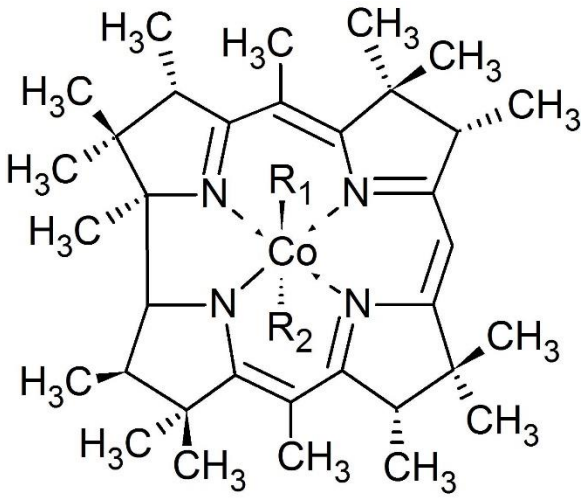


Figure 8.3. Simplified structure used to represent cobalamins. β-axial ligand, R_1, could be a cyano, water, hydroxo or imidazole group. The α-axial ligand, R_2, is DMB in the base-on conformation or CN in the base-off conformation.

SECTION 8.3. RESULTS

Abs spectra of cobalamins. All cobalamins in this chapter exhibit typical Abs spectra (**Fig. G.1.**). An example of a cobalamin with a unique Abs spectrum is MeCbl.[24] The Abs spectra of all cobalamins are unique; each transition is sensitive to the axial ligands (**Table G.1**). The published molar absorptivities were used to determine sample concentration (**Table G.1**). Dicyanocobalamin was prepared by adding excess potassium cyanide to CNCbl (**Fig. 8.4**). The pKa of the DMB ligand in CNCbl is 0.1.[25] The reported maxima for CNCbl are 361, 520, and 551 nm.[26] The γ-maximum is 368 and the α-maximum is 580 nm for diCNCbl.[27]

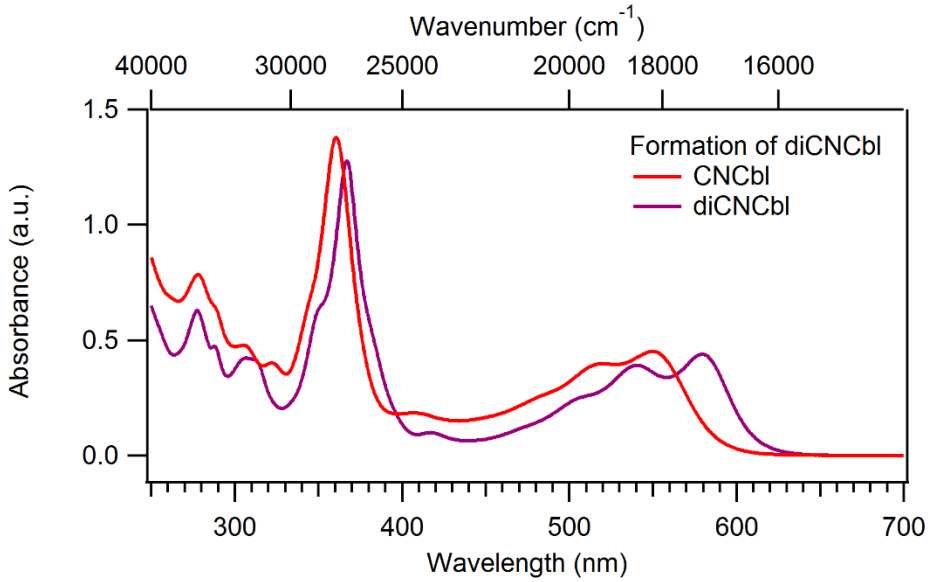


Figure 8.4. Formation of diCNCbl from CNCbl. Cobalamins were dissolved in 50 mM Tris at pH 8.0. Peak maxima for CNCbl: 360.5 (γ), 408 (D/E), 520 (β), and 550 (α) nm. Peak maxima for diCNCbl: 367 (γ), 417.5 (D/E), 541 (β), 580 (α) nm.

ImCbl, bisImCbl, and OCNCbl were synthesized from H_2OCbl. The ligand binding event occurs following the rate-limiting and unimolecular release of the β-axial water ligand.[16] The Abs spectra of HOCbl, ImCbl, and OCNCbl are in the chapter 10 appendix (**Fig. G.2-G.4**). The reported Abs maxima for ImCbl are 357, 413, and 536 nm.[28] Imidazole will bind at the β-axial position; however, bisImCbl will form at high concentrations of imidazole.[28] Upon imidazole binding, all transitions shifted to lower energies as consistent with the literature (**Fig. 8.5**). However, a mixture of species was observed for these cobalamins. A pure sample was required for ^{59}Co NMR; for example, a mixture of HOCbl and H_2OCbl led to no detectible ^{59}Co chemical shift.

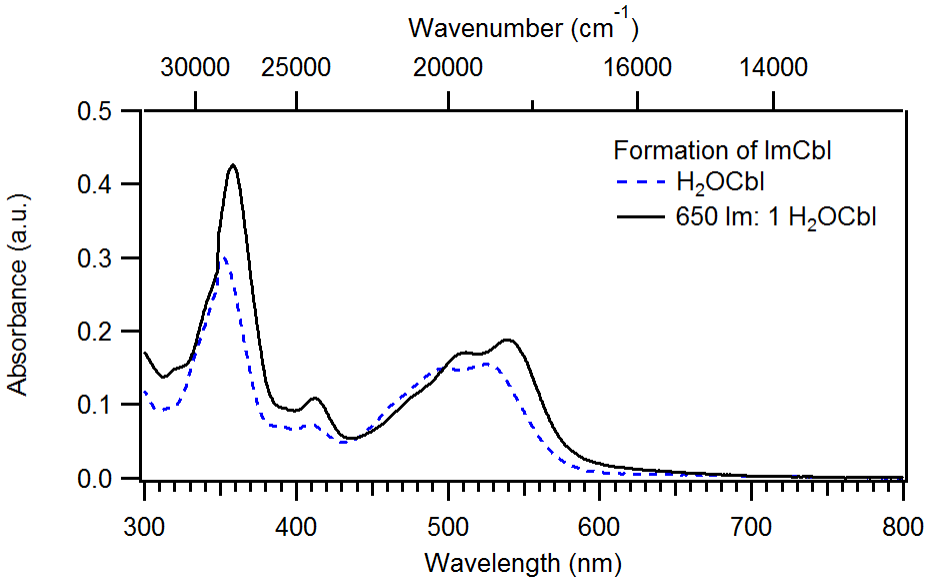


Figure 8.5. Formation of bisImCbl/ImCbl mixture from H_2OCbl in MilliQ water at pH 4.3. Imidazole was added to H_2OCbl in a 650:1 ratio. Sample was diluted 100-fold for Abs spectroscopy. Peak maxima for H_2OCbl: 352 (γ), 412 (D/E), 500 (β), and 528 (α) nm. Peak maxima for bisImCbl/ImCbl: 358 (γ), 414 (D/E), 511 (β), 538 (α) nm.

The stability of the super-reduced Co(I)Cbl was monitored by Abs spectroscopy. The reduction of H_2OCbl to the super-reduced Co(I)Cbl was impeded by ambient oxidation of the cobalamin once removed from the glove box. The cobalamin was removed from the glovebox in a sealed cuvette wrapped in parafilm and monitored by Abs spectroscopy to observe the lifetime of the Co(I)Cbl (**Fig. 8.6**). However, the super-reduced Co(I)Cbl was never observed due to immediate oxidation to Co(II)balamin. The lifetime of Co(I)Cbl was deemed insufficient for the time required to identify the ideal excitation for the ^{59}Co nuclei as H_2OCbl was reformed within 30 min. Therefore, a ^{59}Co NMR spectrum was not collected for this sample.

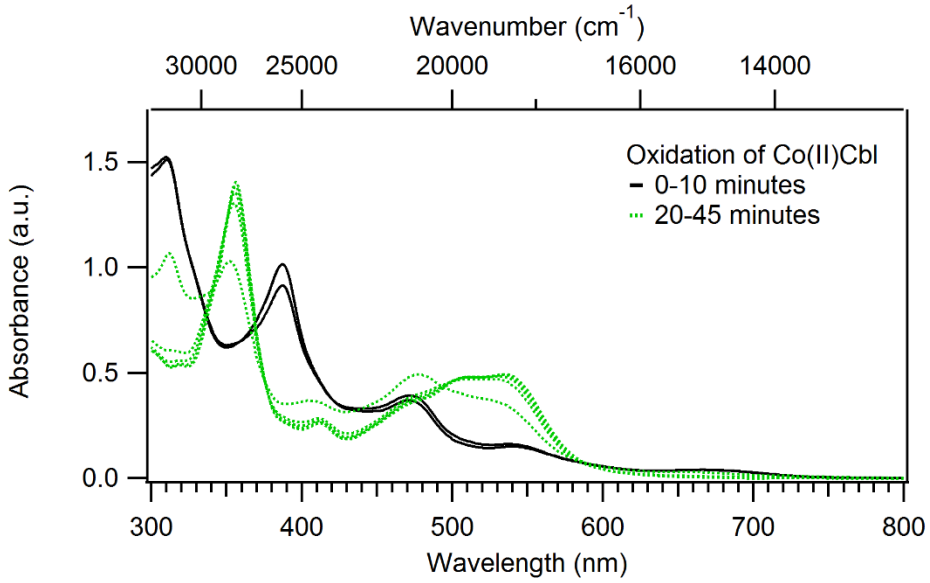


Figure 8.6. The oxidation of Co(II)Cbl to H_2OCbl. The characteristic γ-band at 352 nm of H_2OCbl returned within 25 minutes indicating the Co(III)Cbl species had reformed. No Co(I)balamin was detected.

^{1}H NMR spectroscopy. The aromatic region of the 1D ^{1}H NMR spectra was used to rapidly identify the purity of the synthesized cobalamins.[29] Minor impurities were present in the CNCbl sample but did not impede characterization (**Fig. 8.7**). This spectra was in agreement with previously published data.[29]

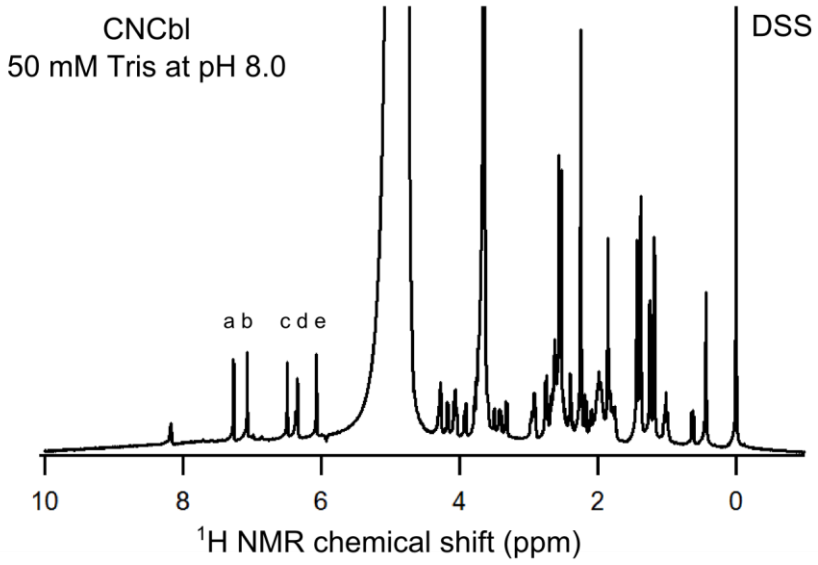


Figure 8.7. ^{1}H NMR spectra of 3 mM CNCbl in 50 mM Tris at physiological pH. The chemical shifts for the labeled aromatic peaks are: (a) 7.28, B2; (b) 7.07, B7; (c) 6.49, B4; (d) 6.36d, R10; (e) 6.07, C10. Sample was internally referenced to DSS.

The 1D ^{1}H NMR spectrum indicated the α-axial DMB ligand of CNCbl had been replaced by a cyano group to form diCNCbl.[15] Following the replacement of DMB by the cyano ligand, the B2 and B4 chemical shifts moved approximately 1 ppm downfield (**Fig. 8.7**). As the environment of the aromatic DMB protons was significantly altered in the base-off conformation, these chemical shifts experienced the greatest change.

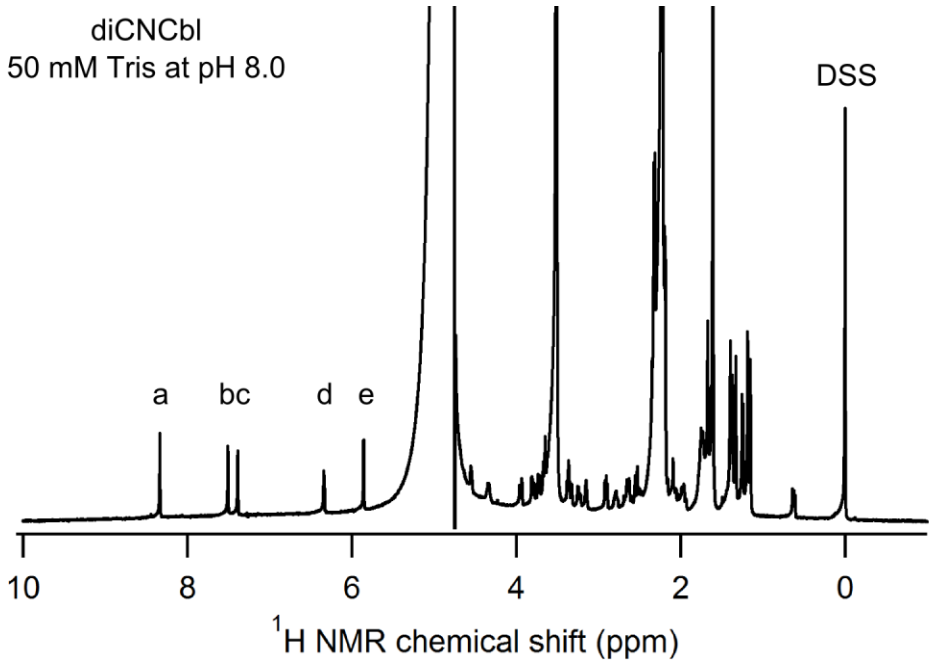


Figure 8.8. ^{1}H NMR spectra of 3 mM diCNCbl in 50 mM Tris at pH 8.0. The chemical shifts for the labeled aromatic peaks are: (a) 8.34, B2; (b) 7.51, B7; (c) 7.39, B4; (d) 6.34d, R1; (e) 5.86, C10.

The aromatic region of the ^{1}H NMR of ImCbl, H_2OCbl, and OCNCbl all showed a mixture of species in aqueous buffer (**Figs. G.5-G.8**). Along with the Abs spectra of these cobalamins differing from the reported literature values, the ^{1}H NMR spectra support a mixture of cobalamins in solution. The ligands (other than cyanide) are not bound tightly to the central cobalt atom and are undergoing exchange with the solvent. This leads to an equilibrium of the starting material and product in solution. Therefore, the collected ^{59}Co NMR spectra of ImCbl and H_2OCbl did not produce a detectable ^{59}Co resonance. ^{59}Co NMR spectra were not collected for OCNCbl because this cobalamin was in equilibrium with the starting material.

^{59}Co NMR spectroscopy. ^{59}Co spectral data was collected of CNCbl in 50 mM Tris at pH 7.5 with 10 % D_2O for locking (**Fig. 8.9**). Prior to scanning a range of frequencies to excite the ^{59}Co nuclei, a simple ^{1}H spectrum was collected and the ^{1}H resonance for water was used to reference. The chemical shift of CNCbl was in agreement with previously reported data.[1a]

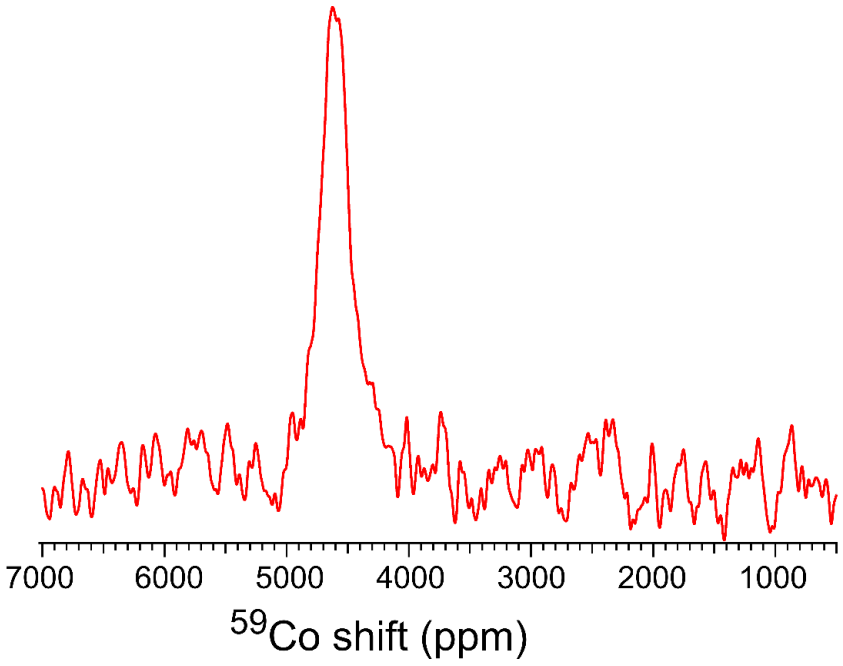


Figure 8.9. ^{59}Co NMR spectra of 3 mM cyanocobalamin in 50 mM Tris with 10 % D_2O at pH 7.5. The chemical shift was 4620 ppm (full width at half height was 25 kHz).

Dicyanocobalamin was synthesized by the addition of excess KCN to a solution containing CNCbl in 50 mM Tris at pH 8. The chemical shift of diCNCbl was upfield of CNCbl reflecting the addition of a second strong field ligand (**Fig. 8.10)**. The second cyanide ligand increases the HOMO-LUMO gap compared to CNCbl leading to a smaller paramagnetic contribution to the chemical shift thus the chemical shift of diCNCbl is upfield relative to CNCbl.

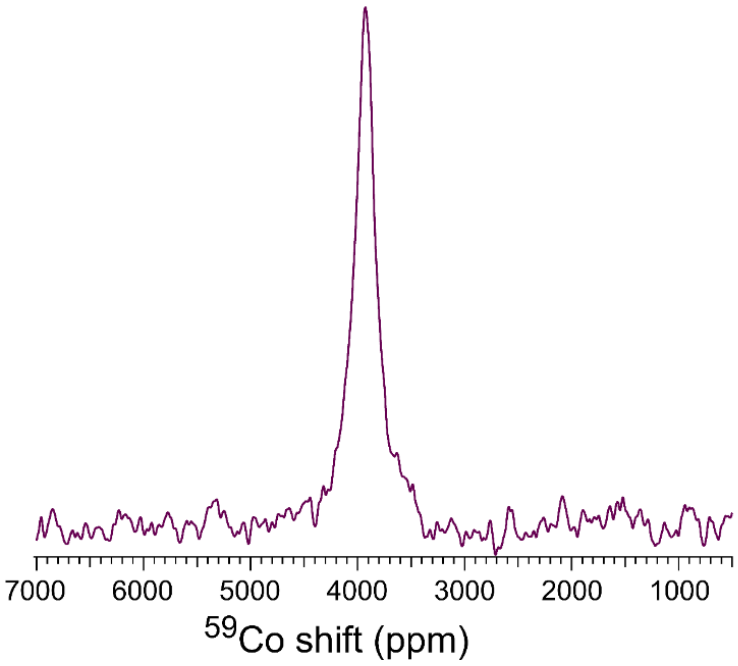


Figure 8.10. ^{59}Co NMR spectra of 4 mM dicyanocobalamin in 50 mM Tris with 10 % D_2O at pH 8. The chemical shift of was 3930 ppm (full width at half height was 24.5 kHz).

Herein lies the main challenge of detecting the ^{59}Co chemical shift of these cobalamins; several equilibria exist leading to a mixture of species changing on the timescale of the NMR experiment. H_2OCbl, HOCbl, and ClCbl were the most likely cobalamins present as mixture of species for both H_2OCbl and HOCbl. The equilibrium of H_2O/Cl exchange at the β-axial position to form ClCbl occurs on a timescale slightly faster than a ^{1}H NMR experiment.[29] Even though it is heavily influenced by pH, the prototopic equilibrium of H_2O/HO results in a mixture of species that are difficult to separate.[29] The mixture of species in equilibrium of a timescale similar to that of the NMR experiment prevents the detection of a cobalt-59 chemical shift of a single cobalamin.

DFT computations. First, the geometry of the simplified cobalamin model was optimized (**Table G.2**). The Löwdin orbitals were plotted to determine the atomic orbital contribution to each molecular orbital. The Co d-orbital contributions were identified by inspecting the contribution from each atomic orbital to the MO. The ligand field splitting is the energy difference between the Co-based HOMO and LUMO (**Table 8.1**).

Table 8.1. Predicted ligand field splitting of cobalamins.

Cobalamin	LFS (eV)
HOCbl	2.92
MeCbl	2.84
diCNCbl	2.72
CNCbl	2.63
AdoCbl	2.60
ImCbl	2.58
H_2OCbl	2.20
Co(I)Cbl	2.14

From these LFS values, the following chemical shifts can be expected: (downfield) Co(I)Cbl > H_2OCbl> ImCbl> AdoCbl> CNCbl> diCNCbl> MeCbl> HOCbl (upfield) (**Fig. 8.10**). The published chemical shifts of the biologically relevant cobalamins are 4650 (CNCbl), 4480 (AdoCbl), and 4215 ppm (MeCbl).[1a]

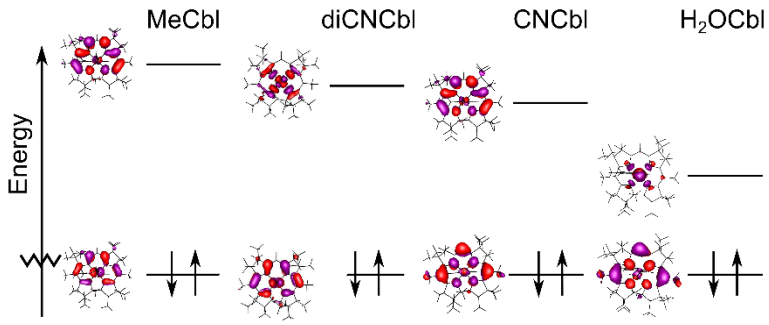


Figure 8.11. Relative LFS of select cobalt tetrapyrroles. A larger LFS should correspond to a more upfield shifted cobalt-59 chemical shift.

SECTION 8.4. DISCUSSION

Large K_d values for ligands lead to a mixture of cobalamin species. Several competing equilibria lead to a mixture of species in solution. For example, the preparation of ImCbl may create eight unique species (**Fig. 8.11**).

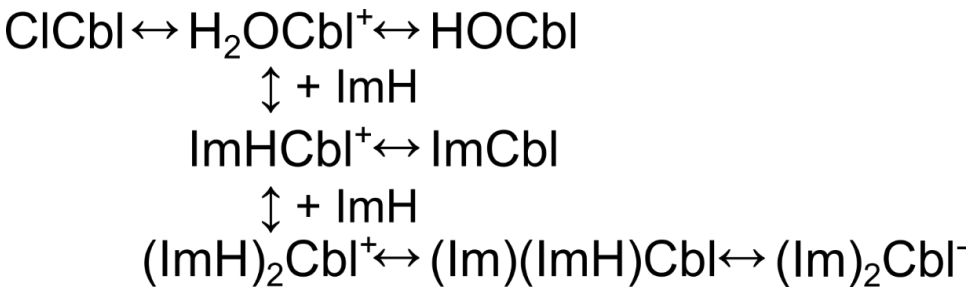


Figure 8.12. Potential cobalamin species in solution upon the addition of imidazole to H_2OCbl. ClCbl is chlorocobalamin, ImH is imidazole, Im^- is imidazolate.

The ClCbl species was proposed to exist in equilibrium that is faster than the NMR timescale; the Cl^- ligand is present due to the buffer conditions.[29] Hydroxocobalamin is also sold as HOCbl·HOAc; there is no evidence of AcOCbl formation.[29] The pKa for the water ligand is 7.8 thus slightly acidic conditions were favored as the HO^- ligand would not be displaced by imidazole.[16] The ideal pH range for imidazole binding is 7-7.8.[16] Acidic conditions would lead to a base-off cobalamin due to the protonation of the DMB ligand (pKa 2.7).[30] The pH-dependent association rate constants were determined at 25 °C with ligand 40-fold molar excess compared to cobalamin.[16] Even with a 8,000,000-fold molar excess of imidazole relative to cobalamin, the solution remains as a mixture of ImCbl and bisImCbl.[28] The pKa of Cbl-bound imidazole is 10; a small fraction of imidazole is present is solution in its anionic imidazolate form.[13] The equilibria of ligand-on/ligand-off for

ImCbl and OCNCbl is supported by their dissociation rates (K_d, M) as compared to CNCbl; K_d (CNCbl)= 10^{-12}, K_d (ImCbl)= 2.6×10^{-5}, and K_d (OCNCbl)= $1.0 \pm 0.3 \times 10^{-3}$.[16] Excluding cyanide, the ligands are not bound tightly to the central cobalt atom and are undergoing exchange with the solvent.

Using aromatic region in ^{1}H NMR spectra to quickly identify cobalamin species in solution. The simple aromatic region of cobalamins was useful for characterizing the cobalamins. Brasch and Finke proposed the aromatic region of the ^{1}H NMR spectra of cobalamins was sufficient for species identification.[29] Typically, five proton resonances are found in the aromatic (specifically 5.50-8.50 ppm) range for cobalamins (**Fig. 8.12**).

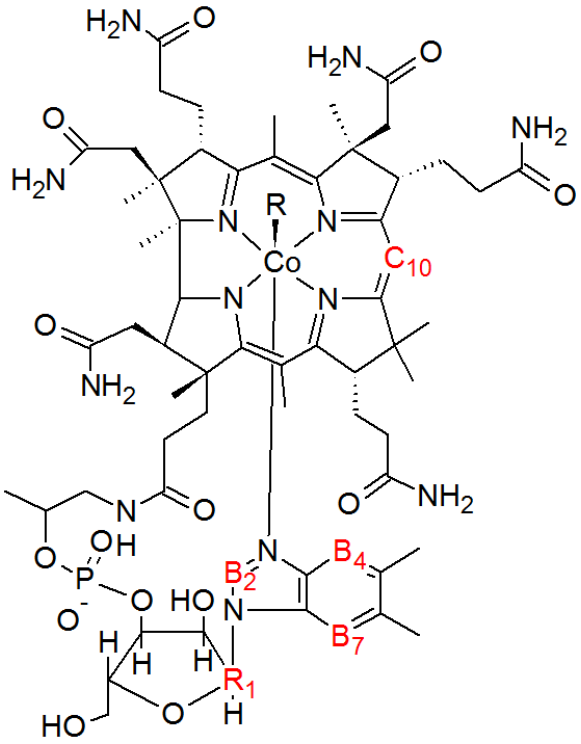


Figure 8.13. The carbon atoms from which the aromatic proton atoms are bound are highlighted in red. In the base-off conformation, the DMB protons (B2, B4, B7) are shifted significantly.

The LFS of the cobalamin relates to the ^{59}Co chemical shift. Computational costs were greatly diminished for cobalamin derivatives by removing the phosphate linker chain connecting the dimethylbenzimidazole ligand to the corrin ring as well as truncating any chains from the corrin ring to methyl groups (**Fig 8.14)**.

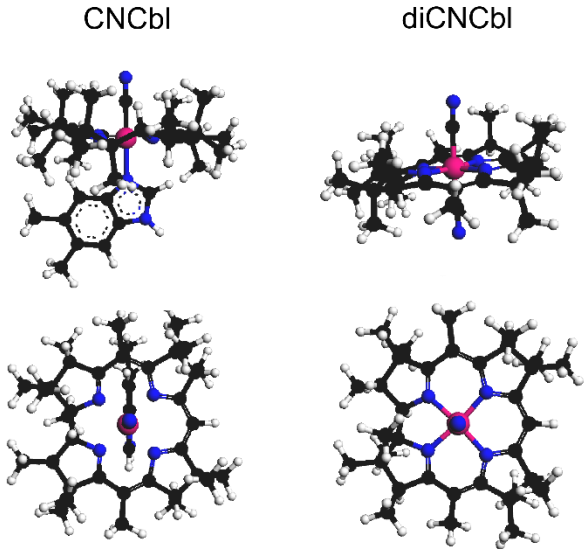


Figure 8.14. Top-down and side view of PBE-DFT optimized structures of CNCbl and diCNCbl. The atoms color scheme of the atoms is as follows: C, black; Co, pink; N, red; O, blue; H, white. The figure was prepared using the program ArgusLab (Planaria software).

The relationship of the ligand field splitting cobalt d-orbitals and chemical shift was determined by DFT calculations. The addition of stronger field ligands will increase the ligand field splitting, The LFS value was derived from the geometry optimized structures of simplified cobalt tetrapyrroles. The LFS was 2.63 eV for CNCbl and 2.72 eV for diCNCbl indicating there should be a greater paramagnetic contribution for CNCbl (**Fig.**

8.15). This prediction was upheld experimentally as the CNCbl chemical shift was located more downfield related to diCNCbl. The smallest ligand field splitting was correctly predicted for Co(I)Cbl. The predicted LFS of MeCbl and AdoCbl were not predicted accurately. Based on the predicted LFS, the chemical shift of AdoCbl would be the most downfield and the most upfield chemical shift would be MeCbl. Experimentally, AdoCbl is less than 200 ppm upfield of CNCbl while MeCbl is about 200 ppm downfield of diCNCbl. The predictive model must be improved to better model these cobalamin model complexes.

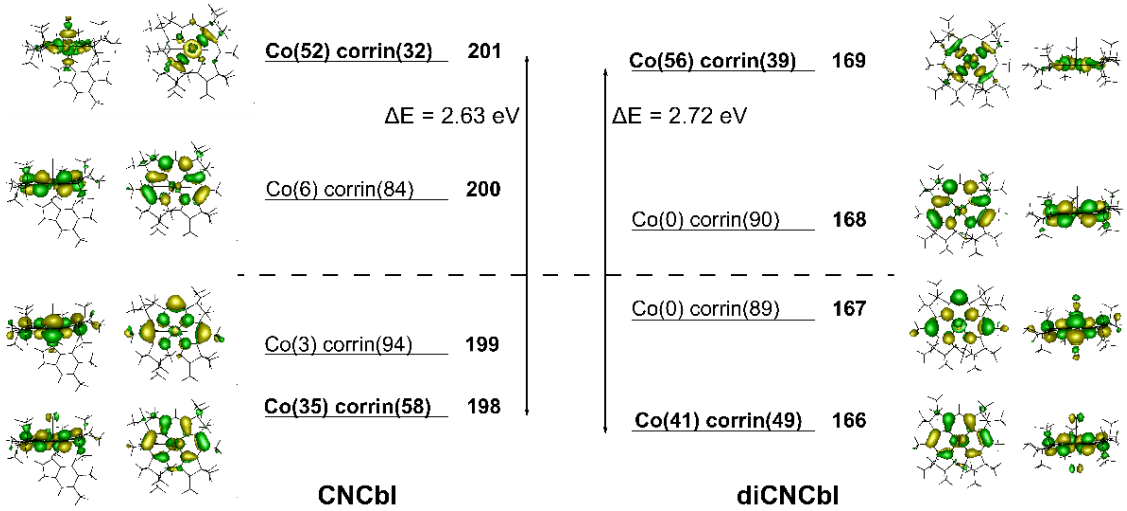


Figure 8.15. Orbital energies calculated by DFT computations for CNCbl and diCNCbl. The percent contribution to the orbital from the cobalt atom and the corrin ring are denoted in parentheses. The outer MO plot is a side view of the simplified cobalamin and the inner MO plot is a top-down view of the cobalamin.

Currently, the model is best suited for determining the chemical shift of an unknown cobalt tetrapyrrole is downfield or upfield as compared to an experimentally verified chemical shift of a known cobalt tetrapyrrole (**Fig. 8.15**). The addition of more species into the model will allow for more accurate predictions of chemical shifts for future cobalt tetrapyrroles. This model will be improved with a different functional or model complex to model the electronic structure of the cobalamins; with $R^2<0.9$, the poorly predicted LFS for MeCbl and AdoCbl prevent any predicted chemical shifts from being accurate (**Fig. 8.16**). The current prediction model erroneously predicted trend of the chemical shifts for the cobalamin species with available ^{59}Co chemical shifts: CNCbl, diCNCbl, AdoCbl, and MeCbl. Specifically, the model was unable to predict that the resonance of MeCbl should be slightly downfield of diCNCbl or that AdoCbl was not the most downfield resonance of the four cobalamins. This is most likely because the GGA-functional (PBE) used to predict the LFS did not accurately model the electronic structure of the simplified cobalamins. New functionals should be used to model these cobalamin model complexes.

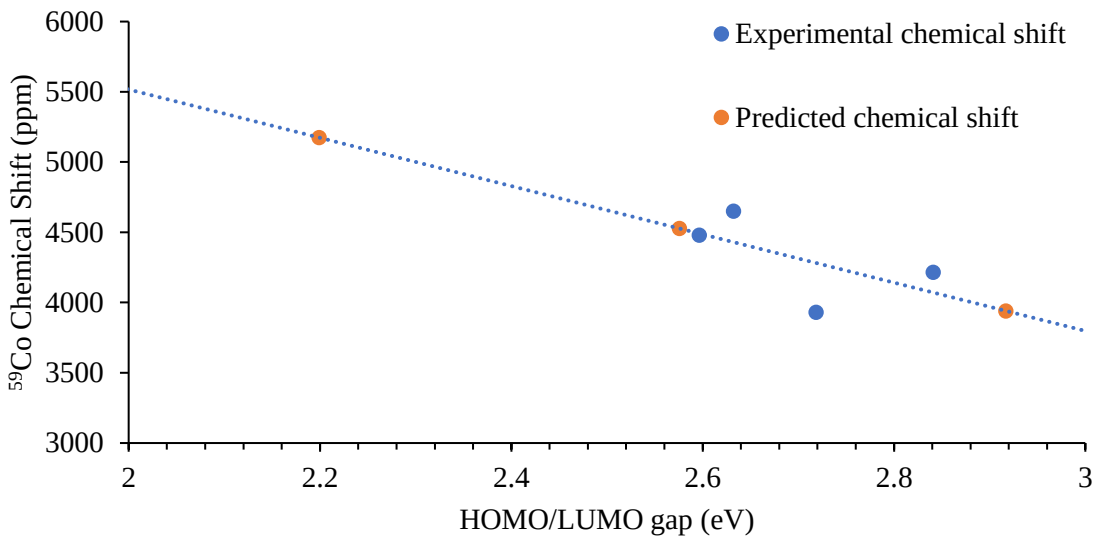


Figure 8.16. Using LFS to predict chemical shift. The blue circles represent the chemical shifts of CNCbl, AdoCbl, MeCbl, and diCNCbl. The orange circles represent H_2OCbl, ImCbl, and HOCbl; their chemical shifts were predicted ($y = -1718.4x + 8953.4$; $R^2 = 0.3518$) from the trend of experimentally derived chemical shifts for CNCbl and diCNCbl. MeCbl and AdoCbl were omitted from the trendline because the current model incorrectly predicted the trend for the LFS. The chemical shift of MeCbl and AdoCbl were taken from reference 1.

SECTION 8.5. CONCLUSIONS AND FUTURE WORK

Improvements to the pulse program for ^{59}Co NMR spectra. A novel pulse program has been developed to collect data on these fast-relaxing tetrapyrroles in less than 30 min. Vitamin B12 and its naturally occurring analogs are some of the most extensively investigated compounds in the chemical and biological literature due to their bio-organometallic nature.[31] Cobalt-59 spectra of some cobalamin samples have been published albeit with much room to improve spectral collection.[1a] Solution-state spectra of cobalamins required 2-4 million scans over days of data collection.[1a] Instead of using antiquated external standard of 1 M aqueous $K_3[Co(CN)_6]$ ($\delta = 0.00$), Ξ-referencing was favored for its facile and fast referencing capabilities.[18-19]

Future cobalt tetrapyrroles. This novel and efficient pulse program can be used on a variety of diamagnetic Co(I) and Co(III) tetrapyrroles. Samples must not contain a mixture of species in equilibrium; it is likely that equilibria of several species led to broadening of the ^{59}Co resonances of HOCbl, H_2OCbl, ImCbl, and bisImCbl beyond detection. The strongly binding cyanide ligand should be used for future cobalt tetrapyrroles. The most facile targets will be dicyanocobalt(III) phthalocyanine and dicyanocobalt(III) protoporphyrin IX (**Fig. 8.16**). The differences in the chemical shifts between the three dicyanocobalt(III) tetrapyrroles will be due to the equatorial ligand.

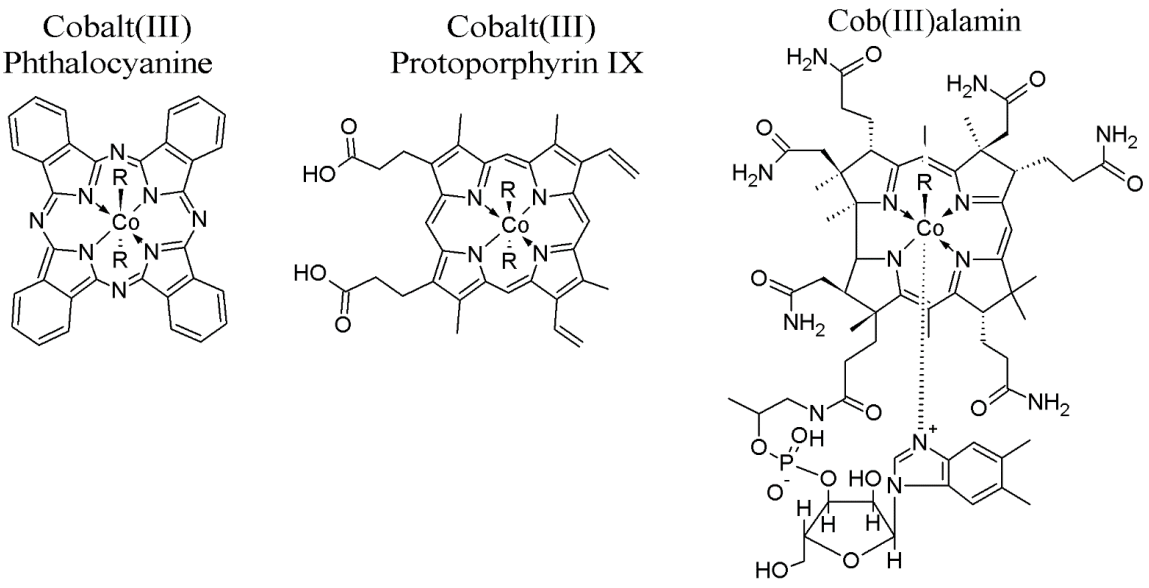


Figure 8.17. General structures of cobalt (III) tetrapyrroles. The axial ligand (R) for cobalt phthalocyanine and protoporphyrin IX will be the same.

Once the library has been expanded to include more Co(III) tetrapyrroles, the data will be paired with the computed LFS values. This predictive model will offer a starting point for probing Co(I) tetrapyrroles. A stronger predictive model is necessary to prevent blind excitation of the ^{59}Co nuclei, the chemical shift range for cobalt complexes is over 20,000 ppm and samples may degrade of the timescale of the experiment (approximately 36 unique experiments over 15 hours of data collection). Once the prediction model has been improved, Co(I) tetrapyrroles will be investigated (**Fig. 8.17**). The reduction of the cobalt tetrapyrroles will follow previously published procedures with zinc dust or titanium (III) citrate.[17, 32] Cuvettes or NMR tube should retain the reductant during data collection to prevent ambient oxidation of the Co(I) species.

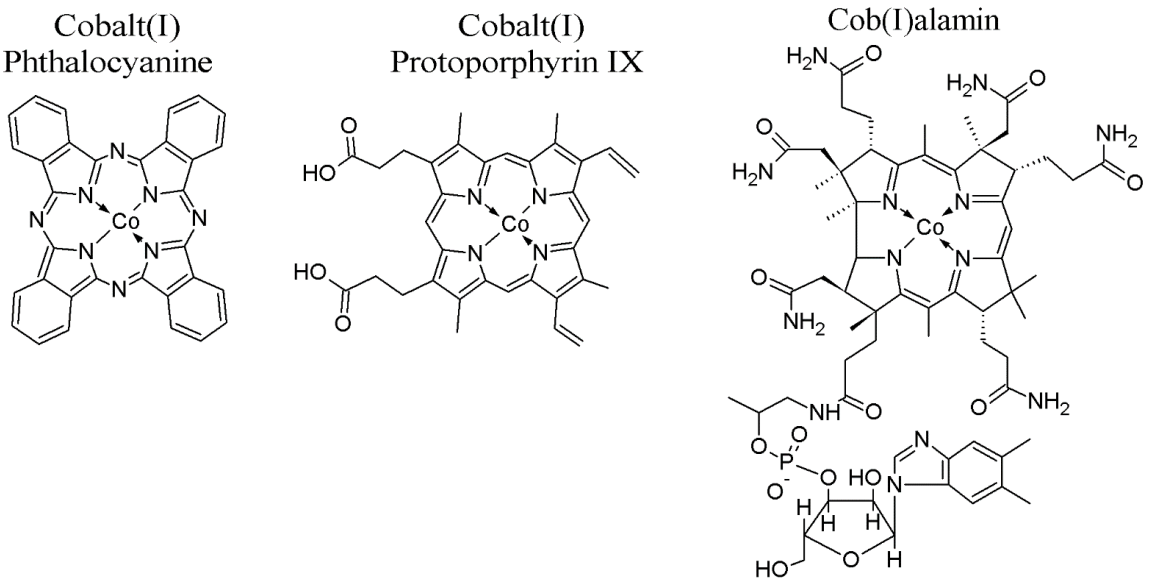


Figure 8.18. Accessible Co(I) tetrapyrroles for ^{59}Co NMR spectroscopy.

The library of cobalt tetrapyrroles could be expanded to include those bound by protein. The ^{59}Co NMR spectra could provide information on the binding site in the protein. The binding of cobalamin in other proteins may be investigated through this technique as sample requirements are compatible with biological samples; the concentration of the cobalamin, and therefore protein, should be 1-3 mM. Several proteins use cobalamin as a catalysis; for example, methionine synthase cycles between Co(I)/Co(III) forms of cobalamin to synthesize the amino acid methionine. The catalytic behavior of methionine synthase could be investigated though ^{59}Co NMR spectroscopy when an inhibitor (such as a cyanide ligand) would be used to prevent catalytic turnover in solution.

Future improvements to DFT modeling of cobalt tetrapyrroles. New functionals should be explored to ensure that PBE is the best fit for the simplified cobalamin. The functionals can be assessed on their modelling capabilities by comparing the TDDFT-predicted Abs spectra of the cobalamin models with experimental spectra as discussed in chapters 2 and 5. A quantum mechanics/molecular mechanics (QM/MM) method may also be worth exploring for more accurate modeling of the cobalamins as it has been successful in the past. Specifically, functionals that are optimized for transition metal complexes or cobalamins should be tested such as BP86 (QM/MM) and B3LYP/G.[33] Two basis sets can be used to include the transition metal and bound atoms in a unique basis set better fit for the metal center.[33b] There is precedence in the literature for a simplified cobalamin, but the model discussed in this chapter differs slightly.

SECTION 8.6. REFERENCES

1. (a) Medek, A.; Frydman, V.; Frydman, L., *Proc. Natl. Acad. Sci. USA* **1997,** *94*, 14237-14242; (b) Gonzalez, J. C.; Banerjee, R. V.; Huang, S.; Sumner, J. S.; Matthews, R. G., *Biochem.* **1992,** *31*, 6045-6056; (c) Nakazono, T.; Parent, A. R.; Sakai, K., *Chem. Commun.* **2013,** *49*, 6325-6327; (d) Lesage, S.; Brown, S.; Millar, K., *Ground Water Monit. Rem.* **1996,** *16*, 76-85.

2. Kliegman, S.; McNeill, K., *Dalton Trans.* **2008**, 4191-4201.

3. Dempsey, J. L.; Brunschwig, B. S.; Winkler, J. R.; Gray, H. B., *Acc. Chem. Res.* **2009,** *42*, 1995-2004.

4. Schrauzer, G. N., *Acc. Chem. Res.* **1968,** *1*, 97-103.

5. Schrauzer, G. N.; Deutsch, E., *J. Am. Chem. Soc.* **1969,** *91*, 3341-3350.

6. Que, L., *Physical Methods in Bioinorganic Chemistry: Spectroscopy and Magnetism*. University Science Books: Sausalito, Calif., 2000.

7. Mason, J., *Chem. Rev.* **1987,** *87*, 1299-1312.

8. Kadish, K. M.; Smith, K. M.; Guilard, R., *The Porphyrin Handbook*. Academic Press: San Diego, 2000.

9. Parella, T. eNMR: NMR Periodic Table. http://www.bruker-nmr.de/guide/eNMR/chem/NMRnuclei.html (accessed July 28, 2014).

10. Chan, J. C. C.; Au-Yeung, S. C. F., *Annu. Rep. NMR Spectrosc.* **2000,** *41*, 1-54.

11. Kelly, A. E.; Ou, H. D.; Withers, R.; Dotsch, V., *J. Am. Chem. Soc.* **2002,** *124*, 12013-12019.

12. Zheng, D.; Yan, L.; Birke, R. L., *Inorg. Chem.* **2002,** *41*, 2548-2555.

13. Hannibal, L.; Bunge, S. D.; van Eldik, R.; Jacobsen, D. W.; Kratky, C.; Gruber, K.; Brasch, N. E., *Inorg. Chem.* **2007,** *46*, 3613-3618.

14. Eilbeck, W. J.; West, M. S., *J. Am. Chem. Soc., Dalton Trans.* **1976**, 274-278.

15. Mishra, P. K.; Gupta, R. K.; Prabhat C. Goswami, P. C.; Venkatasubramanian, P. N.; Nath, A., *Polyhedron* **1982,** *1*, 321-325.

16. Randall, W. C.; Alberty, R. A., *Biochem.* **1967,** *6*, 1520-1525.

17. Liptak, M. D.; Brunold, T. C., *J. Am. Chem. Soc.* **2006,** *128*, 9144-9156.

18. Harris, R. K.; Becker, E. D.; De Menezes, S. M. C.; Goodfellow, R.; Granger, P., *Pure Appl. Chem.* **2001,** *73*, 1795-1818.

19. Harris, R. K.; Becker, E. D.; De Menezes, S. M. C.; Granger, P.; Hoffman, R. E.; Zilm, K. W., *Pure Appl. Chem.* **2008,** *80*, 59-84.

20. Perdew, J. P.; Burke, K.; Ernzerhof, M., *Phys. Rev. Lett.* **1996,** *77*, 3865-3868.

21. Schäfer, A.; Horn, H.; Ahlrichs, R., *J. Chem. Phys.* **1992,** *97*, 2571-2577.

22. Liptak, M. D.; Fleischhacker, A. S.; Matthews, R. G.; Telser, J.; Brunold, T. C., *J. Phys. Chem. B* **2009,** *113*, 5245-5254.

23. Jensen, K. P., *J. Phys. Chem. B* **2005,** *109*, 10505-10512.

24. Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C., *J. Am. Chem. Soc.* **2003,** *125*, 5897-5914.

25. Schneider, Z.; Stroiński, A., Chemistry of cobalamins and related compounds. In *Comprehensive B12: Chemistry, Biochemistry, Nutrition, Ecology, Medicine*, Walter de Gruyter: Berlin, 1987.

26. Hill, J. A.; Pratt, J. M.; Williams, R. J. P., *J. Chem. Soc.* **1964**, 5149-5153.

27. Rosenthal, H. L., Chapter 4: Vitamin B12. In *The Vitamins: Chemistry, Physiology, Pathology, Methods*, Sebrell, W. H.; Harris, R. S., Eds. Academic Press: New York, 1968.

28. Marques, H. M.; Marsh, J. H.; Mellor, J. R.; Munro, O. Q., *Inorg. Chim. Acta* **1990,** *170*, 259-269.

29. Brasch, N. E.; Finke, R. G., *J. Inorg. Biochem.* **1999,** *73*, 215-219.

30. Fanchiang, Y. T.; Bratt, G. T.; Hogenkamp, H. P., *Proc. Natl. Acad. Sci. USA* **1984,** *81*, 2698-2702.

31. Medek, A.; Frydman, L., *J. Am. Chem. Soc.* **2000,** *122*, 684-691.

32. (a) Zehnder, A. J. B.; Wuhrmann, K., *Science* **1976,** *194*; (b) Schindler, O., *Helv. Chim. Acta* **1951,** *34*, 1356-1361.

33. (a) Conrad, K. S.; Jordan, C. D.; Brown, K. L.; Brunold, T. C., *Inorg. Chem.* **2015,** *54*, 3736-3747; (b) Reig, A. J.; Conrad, K. S.; Brunold, T. C., *Inorg. Chem.* **2012,** *51*, 2867-2879.

CHAPTER 9: CONCLUSION

SECTION 9.1. THE ELECTRONIC STRUCTURES OF HYDRAZONE-BASED FLUOROPHORES AND METALLOPROTEINS' COFACTORS WERE DETERMINED SPECTROSCOPICALLY AND COMPUTATIONALLY

The electronic structure of the hydrazone-based dyes provided a more complete understanding of their photophysical properties. One such property, the color of the emission, is related to the energy gap between the emissive and ground electronic states. The anomalous emission from the high-lying excited states of the boron difluorohydrazone (BODIHY) dyes and derivatives resulted in the bright blue emission of these hydrazone-dyes. The large Stoke shifts of triazolopyridinium and triazoloquinolinium was predicted to be correlated to their large structural changes following excitation into the emissive excited electronic state using density functional theory (DFT) and time-dependent DFT calculations.

The catalytic function of metalloproteins arises from the electronic structure of the active site. For example, the supernucleophilic character of Co(I) tetrapyrroles is related to the occupancy of the $3d_{z^2}$-molecular orbital of the central cobalt nuclei. The unique spectroscopic signature of the iron-sulfur cluster of CotA from *C. difficile* based on the electronic structure was used to characterize the cluster. It was proposed that multimerization at the site of the iron-sulfur cluster may be related to spore coat formation in *C. difficile*.

SECTION 9.2. SUPPRESSION OF KASHA'S RULE

In collaboration with the Aprahamian group (Dartmouth College, Hanover, NH), we determined that the BODIHY dyes are the first examples of fluorophores where Suppression of Kasha's Rule (SOKR) can explain the aggregation induced emission (AIE) behavior in viscous environments and the solid state.[1] The enhanced emission in the aggregate state arises from the prevention of internal conversion to the lowest lying excited state, S_1**.** The internal conversion occurs as the molecule freely rotates in solution but this rotation is restricted in a viscous microenvironment. DFT and TDDFT calculations predicted that the bright emissive state was $S_{>1}$ while relaxation from S_1 was primarily non-radiative. This prediction is contrary to Kasha's Rule where internal conversion and vibrational relaxation to the lowest lying excited state results in radiative relaxation.[2] Circular dichroism, excitation-dependent emission, and fluorescence lifetime spectroscopies provided the experimental data to verify this prediction.

SECTION 9.3. NEW HYDRAZONE DYES ARE LIKELY SOKR LUMINOGENS

New SOKR molecules were designed from the basic BODIHY structure and their photophysical properties were predicted with DFT and TDDFT calculations. Several handles are available to alter the photophysical properties of these new hydrazone dyes. Specifically, hydrazones with heavier halogens, an aluminum-containing heterocycle, and substitutions on the phenyl ring of BODIHY were discussed. The most promising synthetic targets were determined based on the predicted barrier to rotation in the ground and excited states, molar absorptivity, and anomalous emission.

The most attractive hydrazone-based fluorophores are *para*-CO_2H BODIHY, aluminum difluorohydrazone (ALDIHY), and *para*-NO_2 ALDIHY. This new BODIHY was predicted to be brighter than the unsubstituted BODIHY and also emit blue light. A bright blue-emitting fluorophore such as *para*-CO_2H BODIHY is desirable to construct a more efficient white LED.[3] It was predicted that *para*-NO_2 ALDIHY would be brighter than the brightest BODIHY to date, *para*-NO_2 BODIHY.[1] In the solid state, the emission of *para*-NO_2 BODIHY was quenched due to the formation of H-aggregates.[1] The replacement of boron with aluminum in the heterocycle may influence solid state packing so that H-aggregates may not form. These ALDIHY derivatives were consistently predicted to be brighter than their BODIHY counterparts. The ALDIHY family of dyes may be the most promising route to develop new SOKR molecules based on the BODIHY family.

SECTION 9.4. AGGREGATION INDUCED EMISSION OF ANOMALOUS FLUOROPHORES

As anomalous emission was responsible for the enhanced emission of the BODIHY dyes in the aggregate state, it is likely that the aggregation induced emission of other anomalous luminogens can be explained by the SOKR mechanism. Three anomalous fluorophores were investigated to determine the influence solvent viscosity had on their anomalous emission: all *trans*-1,6-diphenyl-1,3,5-hexatriene (all-*trans* DPH), zinc meso-tetraphenylporphine (Zn TPP) and azulene (Az).

Az exhibited negative viscosity dependence as the emission intensity was diminished in high viscosity solvents. Aggregation of many Az monomers in high viscosity solvents may have led to quenching by π-stacking or other intermolecular interactions.[4] Due to the negative viscosity dependence of this luminogen, it could be used as a turn-off sensor to monitor changes in viscosity to the microenvironment.

The viscosity sensitivity for all-*trans* DPH and Zn TPP was unclear due to several competing relaxation pathways. For all-*trans* DPH, the photoisomerization led to a mixture of species in solution. The lowered intensity of the DPH emission may have resulted from the existence of *cis*- isomers (lower quantum yield compared to all-*trans* DPH) instead of ACQ. All-*trans* DPH is irradiated by UV light during the collection of emission spectra; the excitation wavelength is 350-360 nm. The viscosity sensitivity of DPH was not determined in part due to the contribution of *cis*- isomers to the emission spectra of DPH. The photoisomerization and potential degradation of DPH will be monitored through ^{1}H

and ^{13}C NMR and the contribution of the *cis*- isomers will be determined by difference emission spectroscopy.

The viscosity sensitivity could not be determined for Zn TPP because π-stacking led to the irreversible formation of multimers. This static quenching was a competing process with the influence viscosity may have had on the anomalous emission of Zn TPP. To circumvent the static quenching, mixtures of alcohols (methanol, ethanol, and propanol) can be used to understand the influence of microenvironment viscosity on Zn TPP emission while preventing π-stacking.

SECTION 9.5. MODELING THE STOKES SHIFT OF TRIAZOLE-BASED DYES

Several GGA- and hybrid GGA-functionals were used to model the large Stokes shifts of the TOP and TOQ dyes: PBE and BLYP (GGA) and PBE0 and B3LYP (hybrid-GGA).[5] PBE most accurately predicted the UV-vis absorption spectra of the TOP and TOQ dyes. Of four functionals tested, PBE had the greatest absolute average error in predicting the Stokes shift. The hybrid-GGA functional, PBE0, most accurately predicted the Stokes shift of the four dyes. Specifically, only PBE0 correctly predicted that the monomers had larger Stokes shifts compared to their dimer counterparts.

As the Stokes shift arises from the relaxation of the molecule in the excited state prior to radiative relaxation, the ground state and excited state geometries predicted by PBE0 were examined. The predicted excited state geometry for the monomers were significantly different from the ground state geometry while the geometry of the dimer did not change following excitation. This difference in changes to the geometry is reflected in the large Stokes shifts. However, of the four functionals tested, a single functional could not accurately model both the Abs spectra and Stokes shifts of the four triazole dyes. New functionals will be tested to see if a different treatment of the exchange and correlation energies may lead to a more accurate prediction of the electronic structure of the triazole-based dyes.

SECTION 9.6. THE MULTIMERIZATION OF AN IRON-SULFUR CLUSTER PROTEIN FROM *CLOSTRIDIUM DIFFICILE* MAY BE RELATED TO THE FORMATION OF THE ENDOSPORE COAT

Several proteins of the Cot family were identified as potential spore coat proteins.[6] In *cotA* knockout strains of *C. difficile*, the endospore coat was malformed and the bacteria was susceptible to cleaning methods such as ethanol.[7] A preliminary UV-vis absorption (Abs) spectrum suggested that CotA contained an iron-sulfur cluster (ISC).[8] The protein multimerized in solution and was isolated in a mixture predominantly containing the CotA dimer. Magnetic circular dichroism (MCD) spectroscopy was used to characterize the putative ISC in CotA. MCD spectra of ISCs are characteristic of unique conformations and oxidations state of the cluster.[9]

The MCD spectra of CotA did not match a single spectrum from the literature of ISCs or other common metal cofactors and likely arose from a mixture of ISCs. The proposed clusters present in CotA are a Rieske [2Fe-2S] and S=3/2 $[4Fe\text{-}4S]^{+}$ cluster. As multimerization of the protein occurred in solution, two Rieske clusters could come together at the protein-protein interface and form the [4Fe-4S] cluster.[10] The most common type of cluster interconversion is two [2Fe-2S] combining to form one [4Fe-4S] cluster.[11] To determine if CotA is involved in the formation of the *C. difficile* spore coat, the morphotype of other strains of *C. difficile* should be characterized.

SECTION 9.7. EXPRESSION OF CBA1 FROM *PHAEODACTYLUM TRICORNUTUM* WAS INSUFFICIENT FOR SPECTROSCOPIC INVESTIGATION FOR COBALAMIN BINDING

CBA1 was expressed in *Escherichia coli* with a tobacco etch virus (TEV) cleavage site. Optimized expression of CBA1 in *E. coli* was through L-arabinose induction.[12] However, the expression of CBA1 was very poor in *E. coli*. A bacterial host, like *E. coli*, may not be fit to express this algal protein. Cobalamin binding by CBA1 could not be determined due to low expression of CBA1 by *E. coli*.

SECTION 9.8. SOLUTION STATE COBALT-59 NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY WAS OPTIMIZED FOR COBALT TETRAPYRROLES

Cobalt-59 NMR spectra were collected for cyanocobalamin (CNCbl) and dicyanocobalamin (diCNCbl) in Tris buffer at physiological pH. The addition of a second cyanide ligand should increase the ligand field splitting and result is a relatively upfield chemical shift. The PBE-predicted LFSs were 2.63 and 2.73 eV for CNCbl and diCNCbl respectively. The chemical shifts for CNCbl matched the previously published reported data (4620 ppm) and the chemical shift of diCNCbl was approximately 600 ppm upfield (3930 ppm).[13] This PBE-based model fell short in predicting the chemical shifts for other biologically relevant cobalamins and needs to be improved.

This optimized pulse program for ^{59}Co NMR spectroscopy can be used for other cobalt tetrapyrroles to understand the role of the equatorial ligand in nucleophilicity. The equatorial ligand will influence the electron density on the cobalt atom.[14] The Co(I) tetrapyrroles will have a relatively smaller LFS and their chemical shifts will be upfield relative to their Co(III) counterparts. The next species that should be characterized by ^{59}Co NMR are dicyano Co(III) phthalocyanine and protoporphyrin IX as well as the super-reduced Co(I) cobalamin, phthalocyanine and protoporphyrin IX species; the computational methodology to predict the ^{59}Co chemical shift should be improved as well.

SECTION 9.9. CONCLUSIONS AND FUTURE DIRECTIONS

The photophysical phenomenon behind the AIE of the BODIHY family of dyes was described by the SOKR mechanism. The SOKR mechanism likely describes the anomalous emission of other AIE fluorophore and two routes to identify new SOKR fluorophores were explored: modeling new hydrazone derivatives with DFT and TDDFT and anomalous fluorophores as identified by the peer-reviewed articles. The three most promising hydrazone derivatives to be synthesized by our collaborators (Aprahamian group, Dartmouth College, Dartmouth NH) as predicted by DFT and TDDFT are *para*-carboxyl BODIHY, ALDIHY, and para-nitro ALDIHY. New anomalous AIE fluorophores that the SOKR mechanism can describe the photophysical properties of may be limited to molecular rotors; this should be the next class of fluorophores should be explored as potential SOKR fluorophores. The large Stokes shift of the triazole-based dyes was modeled by the PBE0 functional but this functional poorly predicted the excited state electronic structure of the dyes at the ground state geometry. More complex functionals will be used to model the Stokes shift and Abs spectra of these triazole-based dyes to see if a single functional accurately predicts both parameters.

The electronic structure of the multimeric protein CotA from C. difficile was used to characterize the cluster as a mixture of Rieske [2Fe-2S] and S=3/2 $[4Fe\text{-}4S]^{+}$ clusters. To determine if the multimerization of CotA is related to spore coat formation, the morphotype of new strains of *C. difficile* should be compared to their corresponding *cotA* knockout strains. The cobalt tetrapyrroles to be synthesized and characterized by ^{59}Co NMR spectroscopy should be the super-reduced Co(I) cobalamin, phthalocyanine and

protoporphyrin IX species as well as dicyano Co(III) phthalocyanine and protoporphyrin IX. New functionals should be tested to improve the computational methodology to predict the ^{59}Co chemical shift.

SECTION 9.10. CHAPTER 11 REFERENCES

1. Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., *Nat. Chem.* **2017,** *9*, 83-87.

2. Kasha, M., *Disc. Faraday Soc.* **1950,** *9*, 14-19.

3. (a) Martin, A.; Long, C.; Forster, R. J.; Keyes, T. E., *Chem. Commun. (Camb)* **2012,** *48*, 5617-5619; (b) Hong, M. K.; Ravva, M. K.; Winget, P.; Bredas, J. L., *Chem. Mater.* **2016,** *28*, 5791-5798.

4. Mei, J.; Leung, N. L.; Kwok, R. T.; Lam, J. W.; Tang, B. Z., *Chem. Rev.* **2015,** *115*, 11718-11940.

5. (a) Perdew, J. P.; Burke, K.; Ernzerhof, M., *Phys. Rev. Lett.* **1996,** *77*, 3865-3868; (b) Lee, C.; Yang, W.; Parr, R. G., *Phys. Rev. B* **1987,** *37*, 785-789; (c) Kim, K.; Jordan, K. D., *J. Chem. Phys.* **1994,** *98*, 10089-10094; (d) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., *J. Chem. Phys.* **1994,** *98*, 11623-11627; (e) Perdew, J. P.; Emzerhof, M.; Burke, K., *J. Chem. Phys.* **1996,** *105*, 9982-9985.

6. Permpoonpattana, P.; Tolls, E. H.; Nadem, R.; Tan, S.; Brisson, A.; Cutting, S. M., *J. Bacteriol.* **2011,** *193*, 6461-6470.

7. Permpoonpattana, P.; Phetcharaburanin, J.; Mikelsone, A.; Dembek, M.; Tan, S.; Brisson, M. C.; La Ragione, R.; Brisson, A. R.; Fairweather, N.; Hong, H. A.; Cutting, S. M., *J. Bacteriol.* **2013,** *195*, 1492-1503.

8. Liu, J.; Chakraborty, S.; Hosseinzadeh, P.; Yu, Y.; Tian, S.; Petrik, I.; Bhagi, A.; Lu, Y., *Chem. Rev.* **2014,** *114*, 4366-4469.

9. (a) Stephens, P. J.; Thomson, A. J.; Dunn, J. B. R.; Keiderling, T. A.; Rawlings, J.; Rao, K. K.; Hall, D. O., *Biochem.* **1978,** *17*, 4770-4778; (b) Collins, D. J.; Zhou, H.-C., Iron-sulfur models of protein active sites. In *Encyclopedia of Inorganic Chemistry*, 2nd ed.; Wiley: 2005.

10. Duin, E. C.; Lafferty, M. E.; Crouse, B. R.; Allen, R. M.; Sanyal, I.; Flint, D. H.; Johnson, M. K., *Biochem.* **1997,** *36*, 11811-11820.

11. Fee, J. A.; Findling, K. L.; Yoshida, T.; Hille, R.; Tarr, G. E.; Hearshen, D. O.; Dunham, W. R.; Day, E. P.; Kent, T. A.; Munck, E., *J. Biol. Chem.* **1984,** *259*, 124-133.

12. Khlebnikov, A.; Risa, O.; Skaug, T.; Carrier, T. A.; Keasling, J. D., *J. Bacteriol.* **2000,** *182*, 7029-7034.

13. Medek, A.; Frydman, V.; Frydman, L., *Proc. Natl. Acad. Sci. USA* **1997,** *94*, 14237-14242.

14. Schrauzer, G. N.; Deutsch, E., *J. Am. Chem. Soc.* **1969,** *91*, 3341-3350.

COMPREHENSIVE LIST OF REFERENCES

1. Abdel-Mottaleb, M. S. A.; Loutfy, R. O.; Lapouyade, R., *PhotoChem. Photobiol.* **1989,** *48*, 97-93.
2. Adamo, C.; Scuseria, G. E.; Barone, V., *J. Chem. Phys.* **1999,** *111*, 2889-2899.
3. Allen, B. D.; Benniston, A. C.; Harriman, A.; Rostron, S. A.; Yu, C., *Phys. Chem. Chem. Phys.* **2005,** *7*, 3035-3040.
4. Baines, S. D.; O'Connor, R.; Saxton, K.; Freeman, J.; Wilcox, M. H., *J. Antimicrob. Chemother.* **2009,** *63*.
5. Barnaba, C.; Humphreys, S. C.; Barden, A. O.; Jones, J. P.; Brozik, J. A., *J. Phys. Chem. B* **2016,** *120*, 3038-3047.
6. Becke, A. D., *J. Chem. Phys.* **1993,** *98*, 5648-5652.
7. Beer, M.; Longuet-Higgins, H. C., *J. Chem. Phys.* **1955,** *23*, 1390.
8. Beinert, H.; Holm, R. H.; Munck, E., *Science* **1997,** *277*, 653-9.
9. Bertrand, E. M.; Allen, A. E.; Dupont, C. L.; Norden-Krichmar, T. M.; Bai, J.; Valas, R. E.; Saito, M. A., *Proc. Natl. Acad. Sci. USA* **2012,** *109*, E1762-1771.
10. Birks, J. B., *Photophysics of aromatic molecules*. Wiley-Interscience: London, New York,, 1970; p xiii, 704 pages.
11. Born, M.; Oppenheimer, R., *Ann. Phys.* **1927,** *389*, 457-484.
12. Borriello, S. P., *Ann. Med.* **1990,** *22*, 61-67.
13. Brasch, N. E.; Finke, R. G., *J. Inorg. Biochem.* **1999,** *73*, 215-219.
14. Burke, K.; Wagner, L. O., *International Journal of Quantum Chemistry* **2013,** *113*, 96-101.
15. Carboni, V.; Su, X.; Qian, H.; Aprahamian, I.; Credi, A., *Chem. Photo. Chem.* **2017,** *1*, 222-229.
16. Cavanagh, J.; Fairbrother, W. J.; Palmer III, A. G.; Skelton, N. J.; Rance, M., *Protein NMR Spectroscopy: Principles and Practice*. Second ed.; Elselvier, Inc. : Burlington, MA, 2007.
17. Chai, J. D.; Head-Gordon, M., *J. Chem. Phys.* **2008,** *128*, 1-16.
18. Chan, J. C. C.; Au-Yeung, S. C. F., *Annu. Rep. NMR Spectrosc.* **2000,** *41*, 1-54.
19. Collins, D. J.; Zhou, H.-C., Iron-sulfur models of protein active sites. In *Encyclopedia of Inorganic Chemistry*, 2nd ed.; Wiley: 2005.
20. Conrad, K. S.; Jordan, C. D.; Brown, K. L.; Brunold, T. C., *Inorg. Chem.* **2015,** *54*, 3736-3747.

21. Crack, J. C.; den Hengst, C. D.; Jakimowicz, P.; Subramanian, S.; Johnson, M. K.; Buttner, M. J.; Thomson, A. J.; Le Brun, N. E., *Biochem.* **2009,** *48*, 12252-64.

22. Croft, M. T.; Warren, M. J.; Smith, A. G., *Eukaryotic Cell* **2006,** *5*, 1175-1183.

23. De Martino, A.; Meichenin, A.; Shi, J.; Pan, K.; Bowler, C., *Journal of Phycology* **2007,** *43*, 992–1009.

24. Deakin, L. J.; Clare, S.; Fagan, R. P.; Dawson, L. F.; Pickard, D. J.; West, M. R.; Wren, B. W.; Fairweather, N. F.; Dougan, G.; Lawley, T. D., *Infect. Immun.* **2012,** *80*, 2704-11.

25. Dempsey, J. L.; Brunschwig, B. S.; Winkler, J. R.; Gray, H. B., *Acc. Chem. Res.* **2009,** *42*, 1995-2004.

26. Ding, D.; Goh, C. C.; Feng, G. X.; Zhao, Z. J.; Liu, J.; Liu, R. R.; Tomczak, N.; Geng, J. L.; Tang, B. Z.; Ng, L. G.; Liu, B., *Adv. Mater.* **2013,** *25*, 6083-6088

27. Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Lidwig, M. L., *Science* **1994,** *266*, 1669-1674.

28. Duderstadt, R. E.; Brereton, P. S.; Adams, M. W.; Johnson, M. K., *FEBS Lett.* **1999,** *454*, 21-6.

29. Duin, E. C.; Lafferty, M. E.; Crouse, B. R.; Allen, R. M.; Sanyal, I.; Flint, D. H.; Johnson, M. K., *Biochem.* **1997,** *36*, 11811-20.

30. Eilbeck, W. J.; West, M. S., *J. Am. Chem. Soc, Dalton Trans.* **1976**, 274-278.

31. Eyring, H., *J. Chem. Phys.* **1937,** *3*, 107-115

32. Fabbiani, F. P. A.; Buth, G.; Dittrich, B.; Sowa, H., *Crystengcomm.* **2010,** *12*, 2541-2550.

33. Fanchiang, Y. T.; Bratt, G. T.; Hogenkamp, H. P., *Proc. Natl. Acad. Sci. USA* **1984,** *81*, 2698-2702.

34. Fee, J. A.; Findling, K. L.; Yoshida, T.; Hille, R.; Tarr, G. E.; Hearshen, D. O.; Dunham, W. R.; Day, E. P.; Kent, T. A.; Munck, E., *J. Biol. Chem.* **1984,** *259*, 124-33.

35. Fiedler, A. T.; Bryngelson, P. A.; Maroney, M. J.; Brunold, T. C., *J. Am. Chem. Soc.* **2005,** *127*, 5449-62.

36. Firth, R. A.; Hill, H. A. O.; Pratt, J. M., *Biochem.* **1967,** *6*, 2178-2189.

37. Förster, T.; Kasper, K., *Z Phys. Chem.* **1954,** *1*, 275-277.

38. Gerding, D. N.; Muto, C. A.; Owens, R. C., Jr., *Clin. Infect. Dis.* **2008,** *46 Suppl 1*, S43-9.

39. Gonzalez, J. C.; Banerjee, R. V.; Huang, S.; Sumner, J. S.; Matthews, R. G., *Biochem.* **1992,** *31*, 6045-6056

40. Guo, B.; Cai, X. L.; Xu, S. D.; Fateminia, S. M. A.; Liu, J.; Liang, J.; Feng, G. X.; Wu, W. B.; Liu, B., *J. Mater. Chem. B* **2016,** *4*, 4690-4695.

41. Guo, Y. X.; Feng, X.; Han, T. Y.; Wang, S.; Lin, Z. G.; Dong, Y. P.; Wang, B., *J. Am. Chem. Soc.* **2014,** *136*, 15485-15488.

42. Haidekker, M. A.; Brady, T. P.; Lichlyter, D.; Theodorakis, E. A., *Bioorg. Chem.* **2005,** *33*, 415-425.

43. Haidekker, M. A.; Theodorakis, E. A., *J. Biol. Eng.* **2010,** *4*, 11.

44. Haidekker, M. A.; Theodorakis, E. A., *J. Mater. Chem. C* **2016,** *4*, 2707-2718.

45. Hamilton, C. L.; Scott, R. A.; Johnson, M. K., *J. Biol Chem.* **1989,** *264*, 11605-13.

46. Hannibal, L.; Bunge, S. D.; van Eldik, R.; Jacobsen, D. W.; Kratky, C.; Gruber, K.; Brasch, N. E., *Inorg. Chem.* **2007,** *46*, 3613-3618.

47. Hanson, G.; Berliner, L.; SpringerLink, Metals in Biology Applications of High-Resolution EPR to Metalloenzymes. In *Biological Magnetic Resonance*, [Online] Springer New York,: New York, NY, 2010; pp. XIX, 419p.

48. Hanzelmann, P.; Hernandez, H. L.; Menzel, C.; Garcia-Serres, R.; Huynh, B. H.; Johnson, M. K.; Mendel, R. R.; Schindelin, H., *J. Biol. Chem.* **2004,** *279*, 34721-32.

49. Harris, R. K.; Becker, E. D.; De Menezes, S. M. C.; Goodfellow, R.; Granger, P., *Pure Appl. Chem.* **2001,** *73*, 1795-1818.

50. Harris, R. K.; Becker, E. D.; De Menezes, S. M. C.; Granger, P.; Hoffman, R. E.; Zilm, K. W., *Pure Appl. Chem.* **2008,** *80*, 59-84.

51. Henriques, A. O.; Moran, C. P., Jr., *Annu. Rev. Microbiol.* **2007,** *61*, 555-88.

52. Hernandez-Marin, E.; Seth, M.; Ziegler, T., *Inorg. Chem.* **2009,** *48*, 2880-6.

53. Hill, J. A.; Pratt, J. M.; Williams, R. J. P., *J. Chem. Soc.* **1964**, 5149-5153.

54. Hong, M. K.; Ravva, M. K.; Winget, P.; Bredas, J. L., *Chem. Mater.* **2016,** *28*, 5791-5798.

55. Hong, Y. N., *Methods Appl. Fluores.* **2016,** *4*, 1-17.

56. Hong, Y.; Chen, S.; Leung, C. W.; Lam, J. W.; Tang, B. Z., *Chem. Asian J.* **2013,** *8*, 1806-1812.

57. Hong, Y.; Lam, J. W.; Tang, B. Z., *Chem. Commun.* **2009**, 4332-4353.

58. Hong, Y.; Lam, J. W.; Tang, B. Z., *Chem. Soc. Rev.* **2011,** *40,* 5361-5388.

59. Hong, Y.; Meng, L.; Chen, S.; Leung, C. W.; Da, L. T.; Faisal, M.; Silva, D. A.; Liu, J.; Lam, J. W.; Huang, X.; Tang, B. Z., *J. Am. Chem. Soc.* **2012,** *134,* 1680-1689.

60. Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K., *J. Chem. Phys.* **2001,** *115,* 3540-3544.

61. Itoh, T., *Chem. Rev.* **2012,** *112,* 4541-4568.

62. Itoh, T.; Kohler, B. E., *J. Phys. Chem.* **1986,** *91,* 1760-1764.

63. Jackson, T. A.; Gutman, C. T.; Maliekal, J.; Miller, A. F.; Brunold, T. C., *Inorg. Chem.* **2013,** *52,* 3356-67.

64. Jensen, K. P., *J. Phys. Chem. B* **2005,** *109,* 10505-10512.

65. Johansson, F. B.; Bond, A. D.; Nielsen, U. G.; Moubaraki, B.; Murray, K. S.; Berry, K. J.; Larrabee, J. A.; McKenzie, C. J., *Inorg. Chem.* **2008,** *47,* 5079-92.

66. Johnson, C. E.; Bray, R. C.; Cammack, R.; Hall, D. O., *Proc. Natl. Acad. Sci. USA* **1969,** *63,* 1234-8.

67. Kaczka, E. A.; Wolf, D. E.; Kuehl Jr., F. A.; Folkers, K., *J. Am. Chem. Soc.* **1951,** *73,* 3569-3573.

68. Kadish, K. M.; Smith, K. M.; Guilard, R., *The Porphyrin Handbook*. Academic Press: San Diego, 2000.

69. Karolczak, J.; Kowalska, D.; Lukaszewicz, A.; Maciejewski, A.; Steer, R. P., *J. Phys. Chem. A* **2004,** *108,* 4570-4575.

70. Kasha, M., *Disc Faraday Soc.* **1950,** *9,* 14-19.

71. Kelly, A. E.; Ou, H. D.; Withers, R.; Dotsch, V., *J. Am. Chem. Soc.* **2002,** *124,* 12013-12019.

72. Kelly, C. P.; LaMont, J. T., *Annu. Rev. Med.* **1998,** *49,* 375-90.

73. Khlebnikov, A.; Risa, O.; Skaug, T.; Carrier, T. A.; Keasling, J. D., *J. Bacteriol.* **2000,** *182,* 7029-7034.

74. Kim, K.; Jordan, K. D., *J. Chem. Phys.* **1994,** *98,* 10089-10094.

75. Kliegman, S.; McNeill, K., *Dalton Trans.* **2008**, 4191-4201.

76. Kottas, G. S.; Clarke, L. I.; Horinek, D.; Michl, J., *Chem. Rev.* **2005,** *105,* 1281-376.

77. Koziar, J. C.; Cowan, D. O., *Acc Chem. Res.* **1978,** *11,* 334-341.

78. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. 2nd ed.; Plenum Publishers: New York, NY, 1999.

79. Larrabee, J. A.; Alessi, C. M.; Asiedu, E. T.; Cook, J. O.; Hoerning, K. R.; Klingler, L. J.; Okin, G. S.; Santee, S. G.; Volkert, T. L., *J. Am. Chem. Soc.* **1997,** *119*, 4182-4196.

80. Lee, C.; Yang, W.; Parr, R. G., *Phys. Rev. B* **1987,** *37*, 785-789.

81. Lesage, S.; Brown, S.; Millar, K., *Ground Water Monit. Rem.* **1996,** *16*, 76-85.

82. Li, Y.; Xu, L. R.; Su, B., *Chem. Commun.* **2012,** *48*, 4109-4111.

83. Liptak, M. D.; Brunold, T. C., *J. Am. Chem. Soc.* **2006,** *128*, 9144-9156.

84. Liptak, M. D.; Fleischhacker, A. S.; Matthews, R. G.; Telser, J.; Brunold, T. C., *J. Phys. Chem. B* **2009,** *113*, 5245-5254.

85. Liu, J.; Chakraborty, S.; Hosseinzadeh, P.; Yu, Y.; Tian, S.; Petrik, I.; Bhagi, A.; Lu, Y., *Chem. Rev.* **2014,** *114*, 4366-469.

86. Liu, J.; Su, H.; Meng, L.; Yihua, Z.; Chunmei, D.; Ng, J. C. Y.; Lu, P.; Faisal, M.; Lam, J. W. Y.; Huang, X.; Wu, H.; Wong, K. S.; Tang, B. Z., *Chem. Sci.* **2012,** *3*, 2737-2747.

87. Liu, Y.; Deng, C.; Tang, L.; Qin, A.; Hu, R.; Sun, J. Z.; Tang, B. Z., *J. Am. Chem. Soc.* **2011,** *133*, 660-663.

88. Lockhart, C. L.; Conger, M. A.; Pittman, D. S.; Liptak, M. D., *J. Biol. Inorg. Chem.* **2015,** *20*, 757-70.

89. Lunde, K.; Zechmeister, L., *J. Am. Chem. Soc.* **1953,** *76*, 2308-2313.

90. Luo, J.; Xie, Z.; Lam, J. W.; Cheng, L.; Chen, H.; Qiu, C.; Kwok, H. S.; Zhan, X.; Liu, Y.; Zhu, D.; Tang, B. Z., *Chem. Commun.* **2001**, 1740-1741.

91. Mapolelo, D. T.; Zhang, B.; Naik, S. G.; Huynh, B. H.; Johnson, M. K., *Biochem.* **2012,** *51*, 8056-70.

92. Marques, H. M.; Marsh, J. H.; Mellor, J. R.; Munro, O. Q., *Inorg. Chim. Acta* **1990,** *170*, 259-269.

93. Martin, A.; Long, C.; Forster, R. J.; Keyes, T. E., *Chem. Commun. (Camb)* **2012,** *48*, 5617-5619.

94. Mason, J., *Chem. Rev.* **1987,** *87*, 1299-1312.

95. McGlynn, S. P.; Daigre, J.; Smith, F. J., *J. Chem. Phys.* **1963,** *39*, 675.

96. Medek, A.; Frydman, L., *J. Am. Chem. Soc.* **2000,** *122*, 684-691.

97. Medek, A.; Frydman, V.; Frydman, L., *Proc. Natl. Acad. Sci. USA* **1997,** *94*, 14237-14242

98. Mei, J.; Hong, Y.; Lam, J. W.; Qin, A.; Tang, Y.; Tang, B. Z., *Adv. Mater.* **2014,** *26*, 5429-5479.

99. Mei, J.; Leung, N. L.; Kwok, R. T.; Lam, J. W.; Tang, B. Z., *Chem. Rev.* **2015,** *115*, 11718-11940.

100. Mesganaw, T.; Im, G. Y.; Garg, N. K., *J. Org. Chem.* **2013,** *78*, 3391-3393.

101. Mishra, P. K.; Gupta, R. K.; Prabhat C. Goswami, P. C.; Venkatasubramanian, P. N.; Nath, A., *Polyhedron* **1982,** *1*, 321-325.

102. Munkholm, C.; Parkinson, D. R.; Walt, D. R., *J. Am. Chem. Soc.* **1990,** *112*, 2608-2612.

103. Nakazono, T.; Parent, A. R.; Sakai, K., *Chem. Commun.* **2013,** *49*, 6325-6327

104. Neese, F., *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2*, 73-78.

105. Ollagnier-de-Choudens, S.; Mattioli, T.; Takahashi, Y.; Fontecave, M., *J. Biol. Chem.* **2001,** *276*, 22604-7.

106. Orenstein, M.; Kimel, S.; Speiser, S., *Chem. Phys. Lett.* **1978,** *58*, 582-585.

107. Pandelia, M. E.; Lanz, N. D.; Booker, S. J.; Krebs, C., *Biochim BioPhys. Acta* **2015,** *1853*, 1395-405.

108. Paredes-Sabja, D.; Shen, A.; Sorg, J. A., *Trends Microbiol.* **2014**, 206-416.

109. Parella, T. eNMR: NMR Periodic Table. http://www.bruker-nmr.de/guide/eNMR/chem/NMRnuclei.html (accessed July 28, 2014).

110. Passerini, A.; Lippi, M.; Frasconi, P., *Nucleic Acids Res.* **2011,** *39*, W288-92.

111. Pejchal, R.; Ludwig, M. L., *PLoS Biol.* **2005,** *3*, e31.

112. Perdew, J. P.; Burke, K.; Ernzerhof, M., *Phys. Rev. Lett.* **1996,** *77*, 3865-3868.

113. Perdew, J. P.; Burke, K.; Ernzerhof, M., *Phys. Rev. Lett.* **1996,** *77*, 3865-3868.

114. Perdew, J. P.; Emzerhof, M.; Burke, K., *J. Chem. Phys.* **1996,** *105*, 9982-9985.

115. Permpoonpattana, P.; Phetcharaburanin, J.; Mikelsone, A.; Dembek, M.; Tan, S.; Brisson, M. C.; La Ragione, R.; Brisson, A. R.; Fairweather, N.; Hong, H. A.; Cutting, S. M., *J. Bacteriol* **2013,** *195*, 1492-503.

116. Permpoonpattana, P.; Tolls, E. H.; Nadem, R.; Tan, S.; Brisson, A.; Cutting, S. M., *J. Bacteriol* **2011,** *193*, 6461-70.

117. Petrenko, T.; Kossmann, S.; Neese, F., *J. Chem. Phys.* **2011,** *134*, 054116.

118. Qian, H.; Cousins, M. E.; Horak, E. H.; Wakefield, A.; Liptak, M. D.; Aprahamian, I., *Nat Chem.* **2017,** *9*, 83-87.

119. Que, L., *Physical Methods in Bioinorganic Chemistry : Spectroscopy and Magnetism*. University Science Books: Sausalito, Calif., 2000.

120. Randall, W. C.; Alberty, R. A., *Biochem.* **1967,** *6*, 1520-1525.

121. Reig, A. J.; Conrad, K. S.; Brunold, T. C., *Inorg. Chem.* **2012,** *51*, 2867-2879.

122. Riegler, M.; Sedivy, R.; Pothoulakis, C.; Hamilton, G.; Zacherl, J.; Bischof, G.; Cosentini, E.; Feil, W.; Schiessel, R.; LaMont, J. T., *J. Clin Invest* **1995,** *95*.

123. Rosano, G. L.; Ceccarelli, E. A., *Front Microbiol.* **2014,** *5*, 172.

124. Rosenthal, H. L., Chapter 4: Vitamin B12. In *The Vitamins: Chemistry, Physiology, Pathology, Methods*, Sebrell, W. H.; Harris, R. S., Eds. Academic Press: New York, 1968.

125. Rouault, T. A., *Nat Chem. Biol* **2015,** *11*, 442-5.

126. Rouault, T. A.; Tong, W. H., *Nat Rev. Mol Cell Biol* **2005,** *6*, 345-51.

127. Rubin, M. S.; Bodenstein, L. E.; Kent, K. C., *Dis Colon Rectum* **1995,** *38*.

128. Ruggiero, M. T.; Gooch, J.; Zubieta, J.; Korter, T. M., *J. Phys. Chem. A* **2016,** *120*, 939-947.

129. Rupnik, K.; Lee, C. C.; Wiig, J. A.; Hu, Y.; Ribbe, M. W.; Hales, B. J., *Biochem.* **2014,** *53*, 1108-16.

130. Saltiel , J.; Krishnamoorthy , G.; Huang , Z.; Ko , D.-H.; Wang, S., *J. Phys. Chem. A* **2003,** *107*, 3178-3186.

131. Saltiel, J.; Sears Jr., D. F.; Sun, Y.-P.; Choi, J.-O., *J. Am. Chem. Soc.* **1992,** *114*.

132. Sasaki, S.; Drummen, G. P. C.; Konishi, G., *J. Mater. Chem. C* **2016,** *4*, 2731-2743.

133. Schäfer, A.; Horn, H.; Ahlrichs, R., *J. Chem. Phys.* **1992,** *97*, 2571-2577.

134. Schindler, O., *Helv Chim Acta* **1951,** *34*, 1356-1361.

135. Schneider, Z.; Stroiński, A., Chemistry of cobalamins and related compounds. In *Comprehensive B12: Chemistry, Biochemistry, Nutrition, Ecology, Medicine*, Walter de Gruyter: Berlin, 1987.

136. Schrauzer, G. N., *Acc Chem. Res.* **1968,** *1*, 97-103.

137. Schrauzer, G. N.; Deutsch, E., *J. Am. Chem. Soc.* **1969,** *91*, 3341-3350.

138. Schwabe, T.; Grimme, S., *Phys. Chem. Chem. Phys.* **2006,** *8*, 4398-4401.

139. Shen, A.; Lupardus, P. J.; Morell, M.; Ponder, E. L.; Sadaghiani, A. M.; Garcia, K. C.; Bogyo, M., *PLoS One* **2009,** *4*, e8119.

140. Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F., *J. Phys. Chem. A* **2006,** *110*, 2235-2245.

141. Sivashanmugam, A.; Murray, V.; Cui, C.; Zhang, Y.; Wang, J.; Li, Q., *Protein Sci.* **2009,** *18*, 936-948.

142. Solomon, E. I.; Zhang, Y., *Acc Chem. Res.* **1992,** *25*, 343-352.

143. Staroverov, V. N.; Scuseria, G. E.; Tao, J. M.; Perdew, J. P., *J. Chem. Phys.* **2003,** *119*, 12129-12137.

144. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., *J. Chem. Phys.* **1994,** *98*, 11623-11627.

145. Stephens, P. J.; Thomson, A. J.; Dunn, J. B. R.; Keiderling, T. A.; Rawlings, J.; Rao, K. K.; Hall, D. O., *Biochem.* **1978,** *17*, 4770-4778.

146. Stephens, P. J.; Thomson, A. J.; Keiderling, T. A.; Rawlings, J.; Rao, K. K.; Hall, D. O., *Proc. Natl. Acad. Sci. USA* **1978,** *75*, 5273-5.

147. Stich, T. A.; Brooks, A. J.; Buan, N. R.; Brunold, T. C., *J. Am. Chem. Soc.* **2003,** *125*, 5897-5914.

148. Studier, F. W., *Protein Expr Purif* **2005,** *41*, 207-234.

149. Su, X.; Liptak, M. D.; Aprahamian, I., *Chem. Commun (Camb)* **2013,** *49*, 4160-4162.

150. Subramanian, S.; Duin, E. C.; Fawcett, S. E.; Armstrong, F. A.; Meyer, J.; Johnson, M. K., *J. Am. Chem. Soc.* **2015,** *137*, 4567-80.

151. Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E., *Phys. Rev. Lett.* **2003,** *91*, 146401.

152. Tong, H.; Hong, Y.; Dong, Y.; Haussler, M.; Li, Z.; Lam, J. W.; Dong, Y.; Sung, H. H.; Williams, I. D.; Tang, B. Z., *J. Phys. Chem. B* **2007,** *111*, 11817-11823.

153. Toohey, J. I., *Biochem.* **1965,** *54*, 934-942.

154. Turek, A. M.; Saltiel, J.; Krishna, T. R.; Krishnamoorthy, G., *J. Phys. Chem. A* **2012,** *116*, 5353-5367.

155. Viscidi, R.; Willey, V.; Bartlett, J. G., *Gastroenterology* **1981,** *81*, 5-9.

156. Wagner, B. D.; Tittelbach-Helmrich, D.; Steer, R. P., *J. Phys. Chem.* **1992,** *96*, 7904-7908.

157. Wang, H.; Zhao, E. G.; Lam, J. W. Y.; Tang, B. Z., *Materials Today* **2015,** *18*, 365-377.

158. Wang, J. X.; Chen, Q.; Bian, N.; Yang, F.; Sun, J.; Qi, A. D.; Yan, C. G.; Han, B. H., *Org. Biomol. Chem.* **2011,** *9*, 2219-2226.

159. Wang, Q.; Ma, D., *Chem. Soc. Rev.* **2010,** *39*, 2387-2398.

160. Wingfield, P., *Curr. Protoc. Protein Sci.* **2001,** *Appendix 3*, Appendix 3F.

161. Yang, Y.; Su, X.; Carroll, C. N.; Aprahamian, I., *Chem. Sci* **2012,** *3*, 610-613.

162. Yuan, W. Z.; Tan, Y.; Gong, Y.; Lu, P.; Lam, J. W. Y.; Shen, X. Y.; Feng, C.; Sung, H. H. Y.; Lu, Y.; Williams, I. D., *Adv. Mater.* **2013,** *25*.

163. Yuan, Y. Y.; Feng, G. X.; Qin, W.; Tang, B. Z.; Liu, B., *Chem. Commun.* **2014,** *50*, 8757-8760.

164. Zechmeister, L., *Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes*. Academic Press Inc Publishers: New York, 1962; p 266.

165. Zehnder, A. J. B.; Wuhrmann, K., *Science* **1976,** *194*.

166. Zhai, D.; Xu, W.; Zhang, L.; Chang, Y. T., *Chem. Soc. Rev.* **2014,** *43*, 2402-2411.

167. Zhang, C.; Jin, S.; Yang, K.; Xue, X.; Li, Z.; Jiang, Y.; Chen, W. Q.; Dai, L.; Zou, G.; Liang, X. J., *ACS Appl Mater. Interfaces* **2014,** *6*, 8971-8975.

168. Zhao, D.; Fan, F.; Cheng, J.; Zhang, Y.; Wong, K. S.; Chigrinov, V. G.; Kwok, H. S.; Guo, L.; Tang, B. Z., *Adv. Opt. Mater.* **2015,** *3*.

169. Zhao, E. G.; Hong, Y. N.; Chen, S. J.; Leung, C. W. T.; Chan, C. Y. K.; Kwok, R. T. K.; Lam, J. W. Y.; Tang, B. Z., *Adv. Healthc. Mater.* **2014,** *3*, 88-96.

170. Zhao, N.; Lam, J. W. Y.; Sung, H. H. Y.; Su, H. M.; Williams, I. D.; Wong, K. S.; Tang, B. Z., *Chem. Eur. J.* **2014,** *20*, 133-138.

171. Zhao, N.; Li, M.; Yan, Y. L.; Lam, J. W. Y.; Zhang, Y. L.; Zhao, Y. S.; Wong, K. S.; Tang, B. Z., *J. Mater. Chem. C* **2013,** *1*, 4640-4646.

172. Zheng, D.; Yan, L.; Birke, R. L., *Inorg. Chem.* **2002,** *41*, 2548-2555.

173. Zhou, Y.; Xiao, Y.; Chi, S.; Qian, X., *Org. Lett.* **2008,** *10*, 633-636.

SECTION A. CHAPTER 2 APPENDIX

Table A.1. PBE/TZVP-optimized Cartesian coordinates of compound 1 in ethylene glycol.

C	-1.41511301077816	-2.31993425220630	-0.61520985290866
C	-1.40701406532813	-3.71591821982346	-0.57786743434804
C	-0.30668186337814	-4.40799512648759	-0.06018174939956
C	0.79378783399944	-3.69139689287567	0.42559175385227
C	0.79594804125873	-2.29652056406443	0.40089963056017
C	-0.30764607823497	-1.60689154104559	-0.12742091100569
N	-0.30657691315747	-0.17781120424458	-0.16108798117405
B	-1.64266382931088	0.63554500903871	-0.13810922718959
N	0.88102113022425	0.37095265955445	-0.28087521051144
C	1.01193743260696	1.69165367290568	-0.18598356844828
C	2.33561638901848	2.17947338968799	-0.38392712757338
C	-0.03198100976346	2.61992069750759	0.18957145892931
C	0.20736996699354	3.97324751231466	0.48917093311117
C	-0.83352328068935	4.77590733126256	0.92837131763798
C	-2.11612266524672	4.22490702666712	1.08091360653522
C	-2.30959913813607	2.89378202587991	0.77019880307341
N	-1.29226776014375	2.10953707240212	0.32471903877845
F	-2.19206972228183	0.70854480209141	-1.43346297408202
F	-2.56107923318527	0.08651095558031	0.76033449272071
N	3.41428864308691	2.59863954720146	-0.54711136802023
H	-2.26816379011816	-1.79255790748200	-1.04130607473889
H	-2.26769241335158	-4.26446739921247	-0.96525303821767
H	1.65313113224732	-4.22117413785057	0.84150538035739
H	1.21570576485732	4.37236288800666	0.38084488112805
H	-0.65268937452763	5.82608207030585	1.16007819824412
H	-2.95669733688322	4.82302807144946	1.42899239901776
H	-3.27946805069103	2.40610598628273	0.85662122168218
H	-0.30761737123410	-5.49948283484991	-0.03225029623393
H	1.64278033714699	-1.73369876399611	0.79365379722324

Table A.2. PBE/TZVP-optimized Cartesian coordinates of compound 2 in ethylene glycol.

C	-0.27824266436347	-2.00071071465730	-1.18354541113825
C	-0.26878251607592	-3.40020372413753	-1.28425611228990
C	0.04238720160818	-4.19775964931398	-0.18142177847486
C	0.36197509633216	-3.60677356421950	1.04230924001226
C	0.36400467463269	-2.21137969840956	1.19431483302119
C	0.02902532283477	-1.43158195882018	0.06854678747656
N	0.00628873826224	0.01300696599758	0.18838776886661
B	-1.30455979821881	0.77919236847752	0.52830580217296
N	1.18962179672511	0.58009456090682	0.18453093454578
C	1.26558960701612	1.91180143280286	0.18134580084007
C	2.58544357372804	2.44316538074667	0.22949775162056
C	0.14642050979667	2.81004405091149	-0.02157775848762
C	0.29974533179843	4.17213658072737	-0.33641300478474
C	-0.81528503379633	4.94025849866177	-0.63340318481713
C	-2.08733207713736	4.34677726064062	-0.63049285489502
C	-2.19687026374450	3.00881278039628	-0.30603087812407
N	-1.10606574465472	2.25988434472427	0.00207527834246
C	-0.57969232740758	-1.15009765863232	-2.39007503862664
C	0.73613109969666	-1.58898190727414	2.51483386732776
F	-2.41895069301052	0.19762591977069	-0.07442697549233
F	-1.48103946526972	0.84934472552538	1.92667209614114
N	3.66086744480836	2.89809738117296	0.27847506032461
H	-0.50339242253938	-3.86303169072359	-2.24563410901240
H	0.04383109187987	-5.28586524138529	-0.27840346593553
H	0.61499800599336	-4.23205567525502	1.90183109651088
H	1.29953133495666	4.60583438595548	-0.35752466415674
H	-0.70016624353114	5.99696748791313	-0.87719268208976
H	-2.98368599080142	4.91774194055743	-0.86711780798794
H	-3.15279282558134	2.48850175548454	-0.26832933852855
H	-0.70153676895853	-1.78005690462748	-3.28077815702915
H	0.22927011986900	-0.43062989183985	-2.58882211617585
H	-1.50351798416595	-0.56896356675921	-2.25374738256714
H	0.83870209255961	-2.36271256840714	3.28658556820866
H	-0.01533168329072	-0.85865996803749	2.84660195024867
H	1.69513265204949	-1.05229218787332	2.44357533495349

Table A.3. PBE/TZVP-optimized Cartesian coordinates of compound 3 in ethylene glycol.

C	-1.15487540137761	-2.18455752421559	-0.43261826582634
C	-1.14183065188517	-3.56844053626647	-0.30753610431063
C	0.01302248212493	-4.22585357124055	0.16048286349812
C	1.15615498082637	-3.47349087620021	0.50108159559022
C	1.13889693192105	-2.09098200243304	0.38018560737830
C	-0.01326087192125	-1.43558321605657	-0.09257151459424
N	-0.02797150697680	-0.01968128293998	-0.21974414581090
B	-1.37596109458425	0.77801899881170	-0.27814047927525
N	1.16200402960516	0.53248151325385	-0.35404778190908
C	1.27982611009070	1.85185555999374	-0.37106548229499
C	2.60770795630392	2.33324426509858	-0.56968994402176
C	0.21505795122366	2.80595464178722	-0.13297629993305
C	0.43694166917754	4.18446594062375	0.02192093802386
C	-0.62596194061398	5.02251911413131	0.32458787493662
C	-1.91032034623698	4.48072249637477	0.48179610908697
C	-2.08447497429222	3.12009683471928	0.31642684222738
N	-1.04532003700349	2.30086802477097	0.00848045694957
C	0.02497751633091	-5.64464432696469	0.29059226272595
F	-1.93803094999277	0.68728463938376	-1.56478420096931
F	-2.27566052691423	0.32324923996846	0.68911806567096
N	3.68802566224329	2.74607715594091	-0.73330261275742
N	0.03506751890817	-6.80815535731009	0.39777358371573
H	-2.04780363935572	-1.69470760255125	-0.81684792658007
H	-2.02577255422777	-4.14501294923092	-0.58105148922430
H	2.04940278234656	-3.97587392337913	0.87350959131408
H	2.01536689776187	-1.50687872134088	0.65684452800114
H	1.44672658843242	4.57859927942528	-0.08996383165075
H	-0.45931886723852	6.09363904237337	0.44522690588447
H	-2.76794793285728	5.10594202178875	0.72475005401072
H	-3.05500384881848	2.63685052868370	0.41858053414400

Table A.4. PBE/TZVP-optimized Cartesian coordinates of compound 4 in ethylene glycol.

C	-1.08060960841683	-2.12320216053083	-0.38807331372434
C	-1.06505259507882	-3.51204616974366	-0.39427215670435
C	0.09434035153354	-4.18586860263420	0.00200130706460
C	1.23981343795834	-3.48990361498998	0.40517581723418
C	1.22072452156605	-2.10281016690941	0.41704876071578
C	0.06320849445223	-1.40578013581638	0.01565667180435
N	0.04346334432952	0.01195095131678	0.02344150826710
B	-1.30710311521796	0.80944549239871	-0.01342094513815
N	1.23644107966402	0.57830903636173	-0.00993587063104
C	1.34729633702502	1.89200911713795	0.09635258713447
C	2.68096713493276	2.39110057960361	0.00124482692211
C	0.26778938920690	2.82323094165554	0.36411170450705
C	0.47646622810031	4.18377375676707	0.64040152725027
C	-0.60419738793176	4.99481040637585	0.95639459418904
C	-1.89153754244526	4.44154350693542	1.00488401796381
C	-2.05123384137263	3.09801515401423	0.72258408347176
N	-0.99465257267973	2.30704901212953	0.40284541428080
F	-1.82933058670237	0.82047305276892	-1.31812669846005
F	-2.23206920234447	0.28090013368756	0.89070711295342
N	3.76487678139818	2.81835062816799	-0.07779027634477
N	0.10933916110320	-5.65073219792701	-0.00253086023986
O	1.16105553580447	-6.22943953737159	0.31559937298505
O	-0.92987668914480	-6.24947576411342	-0.32442506985295
H	-1.97663692188317	-1.60249126467655	-0.71996471621570
H	-1.94259097669164	-4.07174842848629	-0.71449174452073
H	2.12899606113030	-4.03391376877998	0.72088099985501
H	2.09831335132236	-1.54780880516535	0.74375725774410
H	1.48868533556813	4.58662120505037	0.61059816127930
H	-0.44841230838618	6.05263505057732	1.17069769923684
H	-2.76318654819596	5.04447980509134	1.25404022705725
H	-3.02376921760376	2.60813262910473	0.73817459991563

Table A.5. PBE/TZVP-optimized Cartesian coordinates of compound 6 in ethylene glycol.

C	-1.17937098727270	-2.08534812067971	-0.46862257235193
C	-1.21006528150464	-3.47192611574317	-0.39838235327842
C	-0.14849995704769	-4.18099210383874	0.19428706880181
C	0.94954713067365	-3.47538358578228	0.71495765774342
C	0.97285049368440	-2.08213051760183	0.64857105066002
C	-0.08054587259994	-1.37607683375005	0.05174248782881
N	-0.04247932913942	0.04819065513879	-0.01979449334498
B	-1.36165921583080	0.88960399117970	-0.06483499597813
N	1.15890765299280	0.56997819437267	-0.10166893246067
C	1.31243970325044	1.89288886853306	-0.01865774726787
C	2.65181866160813	2.35105291831758	-0.16784704508459
C	0.27435780912582	2.84443524818405	0.30047840519149
C	0.52699252124194	4.19844059991838	0.59277770432266
C	-0.51401853853611	5.02678439531895	0.97733325624292
C	-1.81412151761500	4.50326148179755	1.08293314552798
C	-2.02119706906943	3.17289642419164	0.78155717063905
N	-1.00220159968165	2.36149932666754	0.38847361203306
F	-1.85557067346459	0.94560765282208	-1.38383765355295
F	-2.32664859027635	0.37073367702646	0.80305678653952
N	3.74510681475504	2.74619692138621	-0.29234922118635
O	-0.27749745587788	-5.53954099204496	0.21997229275555
C	0.78852517777918	-6.30243192536959	0.81324873532919
H	-2.00153596851845	-1.55668008055684	-0.95033936413443
H	-2.05582510988913	-4.02650446584730	-0.80835631516223
H	1.77954749202929	-3.99690017207133	1.19031976445664
H	1.81634045259477	-1.53411898089013	1.06868841702888
H	1.54694782566495	4.57598125051990	0.52236127152088
H	-0.32169775958908	6.07615885241329	1.20355412591607
H	-2.65597776982739	5.12290527233050	1.38746385065918
H	-3.00357079900693	2.70573683092812	0.83320256250609
H	0.92050719662542	-6.04094013967406	1.87467152499584
H	1.73449468330434	-6.14959267270696	0.27073325733456
H	0.48197076441702	-7.35061503148947	0.72799604476893

Table A.6. PBE/TZVP-optimized Cartesian coordinates of compound 8 in ethylene glycol.

C	-0.76380164621053	-2.07642663593438	-0.68673288043611
C	-0.73333404557649	-3.46444907360672	-0.66799297795968
C	0.25178166007928	-4.17339958894930	0.07488891552145
C	1.21593847217367	-3.39611722126812	0.77750757155323
C	1.17700726483020	-2.00967915691053	0.75460610208600
C	0.19115691790022	-1.32457318019421	0.02181714015969
N	0.17069168370495	0.09577477722123	-0.01220229452932
B	-1.15604897127962	0.87955675807349	-0.30292916953508
N	1.33446660343540	0.68135395722839	0.15096882843001
C	1.39157777120722	2.01213912708884	0.29642081916549
C	2.70529718466361	2.54199603765147	0.41857001879634
C	0.26193175981045	2.89417505983262	0.43344521765451
C	0.37362370090672	4.25009396085539	0.80513133541778
C	-0.76625406828422	5.00646410538091	1.00946921635661
C	-2.03217633201057	4.41061750671089	0.85484287698920
C	-2.10123736278208	3.08547488577928	0.48124355853676
N	-0.98041590337037	2.34279546646486	0.26239643387808
F	-1.36392641298529	0.96461797077443	-1.69572151481436
F	-2.25272140848623	0.28049616226128	0.32386937963781
N	3.78149078163256	2.99040715453070	0.51657448330062
N	0.26607963118310	-5.54498447545911	0.12170531249526
C	1.37012274270944	-6.24412554960746	0.76828890959522
C	-0.65964726643572	-6.31514345020898	-0.69912712178810
H	-1.52320565218384	-1.57534515662034	-1.28714516365585
H	-1.48056301199226	-4.00533495796219	-1.24685132951072
H	1.99374436018666	-3.88347623626691	1.36375603429601
H	1.91902600474135	-1.43950917701766	1.31439227586670
H	1.36512271957383	4.68229639750706	0.93971157500821
H	-0.68230607100496	6.05498775622199	1.29710906693062
H	-2.95066073576467	4.97309586250795	1.01422115700504
H	-3.04654732910767	2.56636595493302	0.33161252394453
H	1.47480929513711	-5.93637616095683	1.82008508025313
H	2.33318233245413	-6.06526595750085	0.25720413820921
H	1.16453854830581	-7.32004990619317	0.75211537090377
H	-1.70309237661566	-6.03026112805563	-0.49369940611394
H	-0.54676428420334	-7.37778302827143	-0.45745472185730
H	-0.47130180534218	-6.18080613503996	-1.78016655279082

Table A.7. PBE0/TZVP-optimized Cartesian coordinates of compound 1 in ethylene glycol.

C	-1.49151513846324	-2.31851106269232	-0.31053713329598
C	-1.48988936901166	-3.71253782933718	-0.36718192767529
C	-0.30602096193255	-4.43274091665469	-0.19361078526423
C	0.88849316663985	-3.74178859234562	0.03847720204404
C	0.90396900214047	-2.35080659209647	0.10018224338399
C	-0.29065964968901	-1.62979984967472	-0.06967514225591
N	-0.28731186692416	-0.19803606049108	-0.02639609486834
B	-1.63951659895657	0.60961973364651	-0.04052779987255
N	0.90062886570630	0.35081905922207	-0.02850517553786
C	1.04520005148150	1.67117457661946	0.06911107457091
C	2.38646874474976	2.15612349169934	0.01974477199780
C	-0.01802225180115	2.61870726926357	0.28158900557784
C	0.19998288424613	3.99557383079852	0.50938544813918
C	-0.87260006913065	4.82852720106948	0.76730715690298
C	-2.17435036885445	4.29481367550593	0.80940546053193
C	-2.34265307661245	2.94534405242749	0.57893011983169
N	-1.29169697498014	2.12677949764790	0.31442151121997
F	-2.20518823850837	0.59080808584216	-1.31558996264036
F	-2.52938649861769	0.14993295818886	0.92871232109708
N	3.46342119343062	2.60737651849584	-0.01637651777982
H	-2.42141963822230	-1.77405450709042	-0.46242026152580
H	-2.42844674804607	-4.23634659687362	-0.55663182019179
H	1.82142282831253	-4.29155727826913	0.17672560060834
H	1.22239253517116	4.37262130195895	0.48144201261827
H	-0.70723610953814	5.89279537107598	0.94255115295873
H	-3.04177371757316	4.92009945532543	1.01464237243211
H	-3.31625034236948	2.45659923110493	0.59349504665427
H	-0.31227802209098	-5.52338486680275	-0.23853810405092
H	1.83115613444393	-1.80979928356440	0.28128832338974

Table A.8. ωB97X/TZVP-optimized Cartesian coordinates of compound 1 in ethylene glycol.

C	-1.49151513846324	-2.31851106269232	-0.31053713329598
C	-1.48988936901166	-3.71253782933718	-0.36718192767529
C	-0.30602096193255	-4.43274091665469	-0.19361078526423
C	0.88849316663985	-3.74178859234562	0.03847720204404
C	0.90396900214047	-2.35080659209647	0.10018224338399
C	-0.29065964968901	-1.62979984967472	-0.06967514225591
N	-0.28731186692416	-0.19803606049108	-0.02639609486834
B	-1.63951659895657	0.60961973364651	-0.04052779987255
N	0.90062886570630	0.35081905922207	-0.02850517553786
C	1.04520005148150	1.67117457661946	0.06911107457091
C	2.38646874474976	2.15612349169934	0.01974477199780
C	-0.01802225180115	2.61870726926357	0.28158900557784
C	0.19998288424613	3.99557383079852	0.50938544813918
C	-0.87260006913065	4.82852720106948	0.76730715690298
C	-2.17435036885445	4.29481367550593	0.80940546053193
C	-2.34265307661245	2.94534405242749	0.57893011983169
N	-1.29169697498014	2.12677949764790	0.31442151121997
F	-2.20518823850837	0.59080808584216	-1.31558996264036
F	-2.52938649861769	0.14993295818886	0.92871232109708
N	3.46342119343062	2.60737651849584	-0.01637651777982
H	-2.42141963822230	-1.77405450709042	-0.46242026152580
H	-2.42844674804607	-4.23634659687362	-0.55663182019179
H	1.82142282831253	-4.29155727826913	0.17672560060834
H	1.22239253517116	4.37262130195895	0.48144201261827
H	-0.70723610953814	5.89279537107598	0.94255115295873
H	-3.04177371757316	4.92009945532543	1.01464237243211
H	-3.31625034236948	2.45659923110493	0.59349504665427
H	-0.31227802209098	-5.52338486680275	-0.23853810405092
H	1.83115613444393	-1.80979928356440	0.28128832338974

SECTION B. CHAPTER 3 APPENDIX

Table B.1. PBE/TZVP-optimized Cartesian coordinates of dimeta-iPr BODIHY.

47
XYZ file generated by orca_plot on BaseName=m2iPrGS

F	-0.835225	1.071017	-0.877115
F	-1.538067	-0.893227	-1.874713
B	-0.999504	-0.308788	-0.733955
N	-1.850729	-0.623239	0.550174
N	-1.329354	-0.835429	1.732721
N	0.442155	-0.945059	-0.487860
N	0.910750	-1.318395	4.267940
C	-0.011850	-1.007111	1.863656
C	0.922330	-1.145437	0.777083
C	2.262284	-1.558383	0.951383
C	3.066400	-1.770275	-0.152475
C	2.539076	-1.584421	-1.444618
C	1.228971	-1.171899	-1.571213
C	0.465245	-1.177932	3.197052
C	-3.279529	-0.587596	0.517628
C	-3.934723	0.154460	-0.475147
C	-5.332685	0.201024	-0.509730
C	-6.056377	-0.507889	0.460128
C	-5.415796	-1.255541	1.460948
C	-4.019136	-1.291212	1.478617
C	-6.040667	1.026085	-1.574797
H	2.631988	-1.713680	1.965073
H	4.101709	-2.087982	-0.020640
H	3.141559	-1.750043	-2.336367
H	0.749469	-0.996061	-2.533271
H	-3.357940	0.714984	-1.211185
H	-3.489135	-1.863936	2.240662
H	-7.149125	-0.480023	0.433630
H	-5.255112	1.453119	-2.221235
C	-6.212476	-2.018071	2.509426
H	-5.479161	-2.542531	3.145852
C	-7.124292	-3.081658	1.877657
H	-7.898916	-2.620833	1.244936
H	-7.637504	-3.664231	2.657608
H	-6.550130	-3.779163	1.250629
C	-7.013344	-1.065659	3.413755
H	-7.530306	-1.627765	4.206022
H	-6.357156	-0.324253	3.891976
H	-7.776215	-0.517765	2.839555
C	-6.954315	0.160843	-2.458696

H	-7.404354	0.767647	-3.258809
H	-7.775869	-0.277382	-1.871434
H	-6.396365	-0.663478	-2.924952
C	-6.821048	2.197419	-0.953903
H	-7.277418	2.818445	-1.739778
H	-6.163876	2.838166	-0.348336
H	-7.630544	1.834947	-0.301044

Table B.2. PBE/TZVP-optimized Cartesian coordinates of meta-iPr BODIHY.

38
XYZ file generated by orca_plot on BaseName=miPrGS

F	-0.907117	1.046825	-0.923275
F	-1.732558	-0.890182	-1.878295
B	-1.119136	-0.324581	-0.765899
N	-1.925505	-0.605723	0.554490
N	-1.369715	-0.884693	1.707194
N	0.306748	-1.017427	-0.586088
N	0.934364	-1.593857	4.130158
C	-0.062007	-1.139979	1.778493
C	0.822802	-1.285335	0.651442
C	2.147850	-1.763218	0.760641
C	2.902296	-1.967169	-0.379543
C	2.338882	-1.706807	-1.643066
C	1.044383	-1.234230	-1.705466
C	0.455414	-1.386241	3.084935
C	-3.348426	-0.471681	0.586546
C	-3.996991	0.316846	-0.374410
C	-5.385746	0.439516	-0.320236
C	-6.124476	-0.210817	0.669178
C	-5.483582	-1.006318	1.635369
C	-4.093121	-1.129850	1.578171
H	2.545693	-1.972831	1.753558
H	3.926103	-2.334816	-0.298444
H	2.901443	-1.862553	-2.562251
H	0.539415	-1.001705	-2.642068
H	-3.426460	0.843630	-1.137691
H	-3.565405	-1.742203	2.310791
H	-5.895113	1.058137	-1.061794
H	-7.211295	-0.100659	0.687508
C	-6.271924	-1.726688	2.718632
H	-5.541940	-2.303423	3.311998
C	-7.279395	-2.723916	2.124189
H	-8.052230	-2.208697	1.532857

H	-7.789794	-3.280363	2.924532
H	-6.782154	-3.450560	1.465078
C	-6.965031	-0.734403	3.666658
H	-7.483738	-1.269252	4.476414
H	-6.240983	-0.043711	4.122335
H	-7.713691	-0.130824	3.130760

Table B.3. PBE/TZVP-optimized Cartesian coordinates of para-iPr BODIHY.

38
Coordinates from ORCA-job piPr

F	-0.90162811345770	0.53768911626721	-1.64572267642392
F	-1.69651656246834	-1.63116654345684	-1.75502599834796
B	-1.08877429810605	-0.66502391987345	-0.96066135885448
N	-1.88525881886095	-0.41241043821626	0.37140725302572
N	-1.31742171930441	-0.18941195004524	1.53032222541600
N	0.34822879465722	-1.20952471558149	-0.53544252785135
N	1.01488821463207	0.19921304987085	3.99705315356403
C	-0.00257598854266	-0.36517911737523	1.67844438272998
C	0.87896303514747	-0.93915962291248	0.69545484724297
C	2.21560650354128	-1.30618895918802	0.96880479201960
C	2.96625721063020	-1.94393126880343	-0.00094740823994
C	2.38708891068967	-2.23529415324700	-1.25071253309530
C	1.08197644844570	-1.85316649262285	-1.47979721903291
C	0.52722143758582	-0.04850047490466	2.96464310731372
C	-3.30926534550790	-0.30857650820814	0.36621703788435
C	-3.98705314744464	0.00139694393564	-0.82178341849511
C	-5.37732360460076	0.10783694646722	-0.81339017270184
C	-6.12737257699071	-0.08803948442070	0.35502465659719
C	-5.42684115386301	-0.40008493605104	1.53379116013225
C	-4.04086729453470	-0.51388625865925	1.54719031488866
C	-7.64182730150518	0.02748331317317	0.33562585201476
H	2.62490500705975	-1.08319594363594	1.95406835534939
H	3.99949813483565	-2.22606428557215	0.20623371173456
H	2.94678531938986	-2.74257492993161	-2.03514043107808
H	0.56431943207976	-2.03326559864823	-2.42127823774814
H	-3.43282913163421	0.17918984625466	-1.74262790387067
H	-3.50933403887765	-0.75952913503294	2.46614028071881
H	-5.89318728079991	0.35933664636152	-1.74362121471514
H	-5.97447265046882	-0.56966678931947	2.46407076885713
H	-7.92566881268007	0.29968369704889	-0.69568051995523

C	-8.14428340829961	1.14493199410786	1.26481362868348
H	-7.90828080927951	0.92653674566958	2.31766261017021
H	-9.23683662773408	1.25297900086232	1.18369660075013
H	-7.68490400671515	2.11128619149382	1.01201887977520
C	-8.31973514305490	-1.31350428236570	0.66584945418792
H	-9.41421401415737	-1.22297073416186	0.58854207765173
H	-8.08305220659004	-1.63491493498796	1.69169431421562
H	-7.99010342021611	-2.10784038729079	-0.01893433051332

Table B.4. PBE/TZVP-optimized Cartesian coordinates of dimeta-tBu BODIHY.

53
Coordinates from ORCA-job m2tBu

F	-1.02945531216922	1.17005091180303	-0.85293874980512
F	-1.75279996285890	-0.82565486633862	-1.77009756914833
B	-1.19293333805844	-0.20433671446329	-0.65882058342519
N	-2.02414532956613	-0.47621250366823	0.64762885031874
N	-1.48828202916877	-0.68225366233051	1.82400805276769
N	0.25316885691166	-0.83201375055170	-0.41838598802432
N	0.77389357168825	-1.17526385462233	4.33555482203627
C	-0.17124061168476	-0.86963548677166	1.93924615712988
C	0.74818788076548	-1.02152411799040	0.84238055046273
C	2.09073944431665	-1.43169806698899	1.00440961724689
C	2.88269633856218	-1.65036062618865	-0.10677488640256
C	2.34063228518759	-1.47413871579990	-1.39429040803205
C	1.02845611737756	-1.06509780379672	-1.50890791578981
C	0.32045798666428	-1.03429483822258	3.26804452086370
C	-3.45431628492722	-0.41500982887295	0.63652424887705
C	-4.10306554876542	0.33704843651991	-0.35616086181052
C	-5.49662064447660	0.41861102929823	-0.37253193179875
C	-6.21644450025290	-0.26851242492674	0.62469325761145
C	-5.58830333849305	-1.02447296777851	1.62312219555424
C	-4.18792769021045	-1.09449846517153	1.61109448829441
C	-6.37043389035337	-1.78157782405329	2.71013873306141
C	-6.25231679996002	1.24875645232964	-1.42541664500143
H	2.47228606742741	-1.57819544437278	2.01502146066403
H	3.91994603969557	-1.96544682487894	0.01568375414537
H	2.93339969872682	-1.64459026064923	-2.29160257898875
H	0.53809841035763	-0.89739851041678	-2.46697073448521
H	-3.50796601123409	0.87027848209086	-1.09332001947730
H	-3.64535966541923	-1.67153170401737	2.35950557827816

H	-7.30485937669216	-0.21092476123616	0.61451637438276
C	-5.91768847082114	-1.28756130106207	4.10307615266425
H	-6.12453654821408	-0.21390484087098	4.22589146354308
H	-4.84073843364080	-1.44257564684050	4.26040201763468
H	-6.45606523426143	-1.83461011438443	4.89195372341790
C	-6.07889787812054	-3.29360331058927	2.58831839890350
H	-6.62720042936794	-3.84942824185281	3.36442442837293
H	-5.00815533271467	-3.51239710515314	2.71054604893921
H	-6.39182731261265	-3.67773344827842	1.60579179199523
C	-7.88904047402488	-1.57165486101539	2.59683654320853
H	-8.16475331329771	-0.51150118669108	2.70419629032973
H	-8.39683432176334	-2.12796903656511	3.39850821227202
H	-8.28671093138577	-1.93832319140768	1.63840427176931
C	-5.30486883168854	1.87653481948420	-2.46072866946626
H	-4.59210975240263	2.57452466405079	-1.99595173946679
H	-5.89139866854743	2.44532649884194	-3.19760601401191
H	-4.73046985542636	1.11402937741098	-3.00715063190477
C	-7.01999363184158	2.38747555673089	-0.71796487231651
H	-6.32732186099457	3.05368152388254	-0.18229270880966
H	-7.74751279942836	2.00214028575188	0.01150466590229
H	-7.57284374425608	2.98951845550785	-1.45657089514378
C	-7.25589235467308	0.34534182009994	-2.17417824762129
H	-7.79907135486803	0.93226745801565	-2.93130159401144
H	-7.99978525046785	-0.09276619905962	-1.49366998516687
H	-6.73933148657088	-0.47990569193995	-2.68603938853881

Table B.5. PBE/TZVP-optimized Cartesian coordinates of meta-tBu BODIHY.

41
Coordinates from ORCA-job mtBu

F	-1.03734352164282	0.95215284689752	-1.11550055431054
F	-1.71352877139806	-1.13490216476460	-1.84663778176164
B	-1.18123199414182	-0.39895056997454	-0.79395005825637
N	-2.02731294780244	-0.55695325952814	0.52207404612685
N	-1.50400217959164	-0.62395261044311	1.72022394556631
N	0.27020648328484	-0.97731094254522	-0.47268615635771
N	0.73235332133475	-0.80109537943909	4.29597305713245
C	-0.18588870236993	-0.77061714215664	1.87007305751133
C	0.75125211822123	-1.02394069584150	0.80642031433285
C	2.09649332207236	-1.39443401030284	1.02651303151868
C	2.90588097393351	-1.71779745293733	-0.04631200886783

C	2.37849693924175	-1.68776309034477	-1.35116442621536
C	1.06221214167367	-1.31315659241301	-1.52339228306704
C	0.29089612340529	-0.78277916881271	3.21445275376571
C	-3.45709599841895	-0.52913361961559	0.48580750217601
C	-4.11482657003437	0.06841577528229	-0.60102896437249
C	-5.50612224361306	0.08856866895124	-0.60783753180524
C	-6.24093876729709	-0.47291194620140	0.44181422345108
C	-5.59570141287822	-1.07399475436846	1.53439164783639
C	-4.19368884752241	-1.09648924597923	1.53281085239672
C	-6.35906892661614	-1.69893961478774	2.71380050963300
H	2.46726141856507	-1.42722046483720	2.05122854386593
H	3.94576366972666	-2.00205502049751	0.12132325424845
H	2.98516023739494	-1.94355679545636	-2.21838442157415
H	0.58125269838622	-1.25632470330646	-2.49927150367637
H	-3.54882904900954	0.52587847688635	-1.41081011676047
H	-3.64371141229740	-1.55490348641279	2.35431906413522
H	-6.02967065767041	0.55640439027101	-1.44393261377588
H	-7.32985630640988	-0.43868690403971	0.39766125942624
C	-5.99162945487164	-0.94133171466089	4.00994266162459
H	-6.27577252542641	0.11874587188737	3.93823654547875
H	-4.91270913623115	-0.98592360644724	4.21744059850378
H	-6.51903116329579	-1.38073096639881	4.87064175155730
C	-5.96041233833021	-3.18421461054599	2.85754185401856
H	-6.50335020275723	-3.64223661852005	3.69825552108853
H	-4.88443718605938	-3.30152532262602	3.04979740710878
H	-6.20304550876395	-3.74942509089402	1.94494435596664
C	-7.88327707933579	-1.62500357631847	2.52699386776188
H	-8.23801695511222	-0.58599994120643	2.45688376259793
H	-8.38091805232912	-2.08598180676125	3.39308217213572
H	-8.21412860101324	-2.16543591879067	1.62733043483540

Table B.6. PBE/TZVP-optimized Cartesian coordinates of para-tBu BODIHY.

41
Coordinates from ORCA-job ptBu

F	-0.95946094545654	0.76433981629367	-1.36925962152427
F	-1.68454473808367	-1.38894588593971	-1.79254212971633
B	-1.09988474632552	-0.53138009213874	-0.86612394201269
N	-1.89052277691438	-0.51162919533258	0.49155735729790
N	-1.32546509827016	-0.47104816573996	1.67208429640501
N	0.35915388978088	-1.08432588703161	-0.53540953837761

N	1.00667945918999	-0.45210320038971	4.17071249182319
C	-0.00631588094164	-0.63720065196491	1.79004447389911
C	0.88603105477743	-1.01310210480089	0.72436650371611
C	2.23338211715723	-1.38363373773767	0.93211471174229
C	2.99861346699170	-1.82387056269279	-0.13137070665086
C	2.42452921794994	-1.91180534010209	-1.41389914608371
C	1.10798055828749	-1.53431043349185	-1.57529287770260
C	0.52010534650645	-0.53027732570519	3.11148021746336
C	-3.31710520862763	-0.45199059329389	0.50280342062981
C	-4.01051518184541	0.05462368267124	-0.60763001503476
C	-5.40049999940840	0.11316644304160	-0.58389710975996
C	-6.14954277605653	-0.32701976261443	0.52338099243394
C	-5.43073156197211	-0.82889502922339	1.62064200522958
C	-4.03829077251249	-0.89651397773347	1.61759257509414
C	-7.68287416207703	-0.24808227144973	0.49590489217379
C	-8.21664298832899	-1.07457090227913	-0.69585395913878
H	2.64053663721661	-1.32219251885294	1.94164824858767
H	4.03986878323728	-2.10844703546996	0.02679195948297
H	2.99568239183970	-2.26088711849911	-2.27268551875137
H	0.59193489498706	-1.56624758589481	-2.53434141166802
H	-3.46511499482304	0.42245418997224	-1.47606665393694
H	-3.49981629706206	-1.29291295602865	2.47815865391000
H	-5.91087105537489	0.52288369707900	-1.45818619907949
H	-5.95991378332678	-1.18895433415619	2.50268719062902
H	-9.31702057525573	-1.03221883169824	-0.72511542875919
H	-7.91618083608709	-2.12934042051972	-0.60882508533134
H	-7.83953516068138	-0.69543984886891	-1.65636813451963
C	-8.11302286844034	1.22669998487746	0.33039877478108
H	-7.74935740025073	1.83864967078205	1.16951098850939
H	-9.21165284568971	1.30216333506242	0.30214010328837
H	-7.72177868613026	1.66264783566761	-0.59946201094621
C	-8.32034516462040	-0.79549995539944	1.78437143838896
H	-9.41534592183357	-0.71302875760061	1.71469932507782
H	-8.00028621598234	-0.23059724743275	2.67274707937949
H	-8.07627249054293	-1.85681875336426	1.94206639805074

Table B.7. PBE/TZVP-optimized Cartesian coordinates of para-CN BODIHY.

30
Coordinates from ORCA-job pCN

F	-0.34075379822997	-2.18082417875758	-0.60010104022377
F	0.03715494427873	-1.62109561264419	1.61043488270492
B	0.09269456377255	-1.19779337976225	0.28270742225578
N	-0.71161774966164	0.14235479178497	0.10172127213579
N	-0.16874077168757	1.33823980926271	0.14550953313030
N	1.61682453966562	-0.86805041280275	-0.05880001020081
N	2.09597013607708	3.88547738840214	0.32920135481096
C	1.15341089172536	1.47384433121721	0.16325971625692
C	2.10971204858503	0.40389346524374	0.00171933464139
C	3.49277180520877	0.62155088488293	-0.17182453205879
C	4.33288390220592	-0.45164621673341	-0.41022771057928
C	3.79791748933401	-1.74978688115288	-0.48844538416809
C	2.43995541767033	-1.91831192802059	-0.30734861465725
C	1.64247683065754	2.81178524574892	0.25620235829033
C	-2.12991223800897	0.13544852006943	-0.02356736105736
C	-2.86282881286057	-0.99995566563466	0.36382645544342
C	-4.24686460609666	-1.00742919117649	0.23919613728398
C	-4.92366195862086	0.11561581435290	-0.27314464692024
C	-4.18325158181631	1.25220984936676	-0.65571844394832
C	-2.80080654905262	1.26017164970963	-0.53572554764429
C	-6.34459241650719	0.10261190587953	-0.40551703765110
N	-7.50719552460703	0.09212348707540	-0.51460905797048
H	3.87109381123403	1.64264487397248	-0.12202026610578
H	5.40338176636105	-0.28728182539714	-0.54354130287956
H	4.42821857466152	-2.61654038060793	-0.68109702989353
H	1.94928199917217	-2.89019656204249	-0.34769183074316
H	-2.35581618439750	-1.86942447415866	0.77831096295680
H	-4.81520786954144	-1.88585894162176	0.54586006787036
H	-4.70207021133063	2.12322645183709	-1.05739492009049
H	-2.22358704219073	2.13438892470693	-0.83308983898864

Table B.8. PBE/TZVP-optimized Cartesian coordinates of dimeta-CN BODIHY.

31
Coordinates from ORCA-job m2CN

F	-0.32865250487529	-2.18353151511004	-0.62622166356282
F	-0.00208399186396	-1.61244983794687	1.58828841247358
B	0.08920873389470	-1.19807468216780	0.26069312545161
N	-0.70790427923809	0.14046905638150	0.04918055075683
N	-0.16912892091904	1.33994734845171	0.08961354867579
N	1.62092027992399	-0.87018537888194	-0.04470255828001
N	2.09299960832451	3.88897562010580	0.28764796965398
C	1.14992961230949	1.47686157125943	0.13205795806105
C	2.11084603514607	0.40265134761908	0.00755716049600
C	3.49616469871889	0.62007632406660	-0.13702477849801
C	4.34337422310734	-0.45618522301070	-0.33811101910478
C	3.81237895490370	-1.75544525747425	-0.40706645532895
C	2.45017707846893	-1.92352867082466	-0.25555838487048
C	1.63858995822955	2.81572949260383	0.21813006801736
C	-2.12428059321702	0.13217490687135	-0.09837089908299
C	-2.85974018106165	-0.99519726302058	0.29323465011887
C	-4.25568377437761	-1.00150891363684	0.13750045587822
C	-4.92607955700604	0.10814960923410	-0.39889495024857
C	-4.17680735074723	1.23217513543090	-0.78135803394932
C	-2.78150979904549	1.24695063654221	-0.63770915956016
C	-5.00447358280228	-2.15275464035917	0.53989950943835
N	-5.61799044896132	-3.08963477655480	0.86615016415292
C	-4.84276354184323	2.37013578489600	-1.33802385509514
H	3.87250911365379	1.64222870599372	-0.09487823106355
H	5.41648122248618	-0.29207737693606	-0.44890549184402
H	4.44793803038927	-2.62446683913452	-0.57003985863291
H	1.96190703049842	-2.89685479884599	-0.29128734813160
H	-2.36551584655673	-1.86123401525902	0.72905902337698
H	-2.20454628147886	2.11942442805770	-0.93907162763692
H	-6.00864790410579	0.09666672938635	-0.51928536875036
N	-5.38796364395519	3.29570515226295	-1.79269953991097

Table B.9. PBE/TZVP-optimized Cartesian coordinates of meta-CN BODIHY.

30
Coordinates from ORCA-job mLCN

F	-0.33783455031039	-2.18542777126804	-0.61281542616357
F	0.01220035142540	-1.61359548397896	1.59917741141137
B	0.08402461617456	-1.19806744534891	0.26995475318326
N	-0.72000992857543	0.13863762088590	0.07170612349271
N	-0.18470147735804	1.33565868575395	0.12700650138693
N	1.61099573162459	-0.86399225579941	-0.05297398789660
N	2.06589784759885	3.89218829566783	0.34923055802073
C	1.13738400364542	1.47646304602439	0.16710853367363
C	2.09868870061999	0.40971156777188	0.01708264146261
C	3.48357309320556	0.63143933747795	-0.13722719801147
C	4.32986290628935	-0.43872209965607	-0.36659870153171
C	3.80016175449697	-1.73842292630211	-0.45525188950949
C	2.44040042352671	-1.91137003562055	-0.29297704886267
C	1.61948307630094	2.81597804070324	0.26880265771954
C	-2.13903284257685	0.12667832782505	-0.08036850663853
C	-2.87124210646027	-0.99805936285770	0.32078233783894
C	-4.26688446282345	-1.00369939858688	0.16295668683689
C	-4.93123979201303	0.10831983307390	-0.38711585579447
C	-4.18606778108040	1.22093860080487	-0.77595861139122
C	-2.80025923107812	1.23665534298023	-0.63248188676559
C	-5.01914273719207	-2.14908185017776	0.57363143375369
H	3.85845023412745	1.65340311263881	-0.08001507633801
H	5.40162572795442	-0.27077090323122	-0.48471019063134
H	4.43563408940694	-2.60294088387528	-0.64079402086516
H	1.95295306379285	-2.88455217353511	-0.34140179898906
H	-2.37682837014965	-1.85967596242969	0.76524908563663
H	-4.69087209799072	2.08647745425448	-1.20803459372281
H	-2.21520781514241	2.10259295851766	-0.93948285014036
H	-6.01481606967832	0.08766632785576	-0.50560799300240
N	-5.64120034976086	-3.07914112356821	0.90508841283755

Table B.10. PBE/TZVP-optimized Cartesian coordinates of para-F BODIHY.

29
Coordinates from ORCA-job pF

F	-1.08850441757859	0.92192161393652	-1.16172917080849
F	-1.54832372064538	-1.18585044614304	-1.98973817937433
B	-1.11843553674873	-0.44766991002251	-0.89117511728613
N	-1.99355292679237	-0.73179684048218	0.38367189827964
N	-1.51038202869671	-0.84374931612864	1.59626752497949
N	0.36602917867699	-0.90997351760824	-0.54004002289845
N	0.63474921571857	-1.05743660840719	4.24481801821541
C	-0.19263148207895	-0.91281522248293	1.79052131292025
C	0.80040684710045	-1.01438557809787	0.75208282412855
C	2.16211119398955	-1.29233028176611	1.00428971494552
C	3.03627003634975	-1.46729194659085	-0.05201121884584
C	2.55924183456340	-1.37735951549668	-1.37323654802519
C	1.22385668619212	-1.09766223383363	-1.57612666331022
C	0.23368163135055	-0.99086973795686	3.14964708996691
C	-3.41722150051719	-0.78962615049177	0.29496966795976
C	-4.07596983354297	-0.20112762121903	-0.79730728984823
C	-5.46598666939711	-0.25317905887349	-0.88389742203421
C	-6.18171927272220	-0.89465695439647	0.12265232796849
C	-5.55039569843919	-1.48222527307777	1.21676300321508
C	-4.16132235309801	-1.43528246462386	1.29624092406378
F	-7.53587476980299	-0.94898215341122	0.03701934005605
H	2.49453653266990	-1.37376347354741	2.03925811870390
H	4.08864652393977	-1.68153965901692	0.14071442057421
H	3.21775791230537	-1.51616696562610	-2.22923252800009
H	0.77823194480577	-1.00557299622573	-2.56583843968639
H	-3.50778441161398	0.31578953547585	-1.56889274538587
H	-3.64031755673443	-1.88838289963336	2.13868827040161
H	-5.99544982636032	0.20296047232401	-1.72063482325540
H	-6.14446669889306	-1.98145429357650	1.98277512838021

Table B.11. PBE/TZVP-optimized Cartesian coordinates of meta-F BODIHY.

29
Coordinates from ORCA-job mF

F	-1.04900222309872	0.96876497295846	-1.05317256636386
F	-1.59166688824614	-1.10914580422881	-1.90767863691554
B	-1.10106967831726	-0.40485292845108	-0.81364244845378
N	-1.92506410844457	-0.70070565028314	0.49339759794551
N	-1.38947915022845	-0.85380607396944	1.68017368296963
N	0.38900428807378	-0.90671973159716	-0.53803576686671
N	0.86786753186617	-1.16816487001667	4.22411510834546
C	-0.06810411601892	-0.95230287587296	1.81428615923429
C	0.87753833991061	-1.04793735037133	0.73004997095157
C	2.24253726322162	-1.35686620472253	0.91660298522363
C	3.06642703923165	-1.52363608961302	-0.18130227799431
C	2.53392408885410	-1.39505784146182	-1.47718095046483
C	1.19616222619895	-1.08630595122500	-1.61492110270749
C	0.41712112726706	-1.06967627038281	3.15106784603098
C	-3.35291576360265	-0.71799533936352	0.47224061328426
C	-4.04694779445745	-0.11536840071501	-0.58992483811244
C	-5.44187947833475	-0.13458338133696	-0.59573077626302
C	-6.16194217478755	-0.74670687693805	0.43303735637249
C	-5.44224809411871	-1.33227560033352	1.47215394288814
C	-4.05601938249320	-1.34100304771386	1.51536377005736
F	-6.12376798968095	-1.93615966328999	2.48151181987407
H	2.61811963637317	-1.46817386288184	1.93389276549854
H	4.12190330559210	-1.76134182875455	-0.04017608036050
H	3.15149972509574	-1.52649139638765	-2.36429057589282
H	0.70930583982729	-0.96369492991787	-2.58161129281973
H	-3.50406744123909	0.38205980903936	-1.39098093911062
H	-3.52852545470770	-1.80859717341913	2.34484026359722
H	-7.25201353302947	-0.77175298381676	0.44185254215928
H	-5.97862236970667	0.33954001906665	-1.41884479710679

Table B.12. PBE/TZVP-optimized Cartesian coordinates of dimeta-F BODIHY.

29
Coordinates from ORCA-job m2F

F	-1.07049598582124	0.94470177354010	-1.07606023141799
F	-1.59123521021401	-1.14900457903484	-1.90276474386355
B	-1.10651655680391	-0.42568894617245	-0.81888898662626
N	-1.92803535375144	-0.71489683481578	0.49134166837021
N	-1.39253460914009	-0.87706342704829	1.67808685570438
N	0.38941492976156	-0.90801926928307	-0.53950953464092
N	0.86163493593043	-1.22012258190870	4.21999804266856
C	-0.07329002022560	-0.97978469327308	1.81185569094401
C	0.87594247552982	-1.05916444856679	0.72750921747669
C	2.24243684798045	-1.35767335712682	0.91509998895375
C	3.07145941959040	-1.50300675493533	-0.18260915368973
C	2.54186380259486	-1.36247518652242	-1.47784265291763
C	1.20124594813486	-1.06522789778342	-1.61614314299694
C	0.41088826156188	-1.11097037021784	3.14812394212315
C	-3.35489995062407	-0.71942920412507	0.47449532478849
C	-4.03439791083535	-0.12530945359416	-0.60085520353477
C	-5.42378222519600	-0.14476919209348	-0.58621106828569
C	-6.17027492592140	-0.72532173075921	0.43479344823369
C	-5.44985360816977	-1.29913585544186	1.48034516914302
C	-4.06354552658232	-1.32149467288615	1.52648694842202
F	-6.14103811103323	-1.87860956834405	2.49345153848644
F	-6.08642755607699	0.43853449042081	-1.61576902185633
H	2.61608417267019	-1.47745884714804	1.93210862337963
H	4.12870542384719	-1.73259776516064	-0.04129854752477
H	3.16330539364810	-1.47666650354047	-2.36465901381367
H	0.71630645444185	-0.93574337308615	-2.58292087750241
H	-3.50997581616978	0.36333799641885	-1.41884719402329
H	-3.54166560380568	-1.78074742061640	2.36318732981974
H	-7.25930444332070	-0.72646731589522	0.41962893118018

Table B.13. PBE/TZVP-optimized Cartesian coordinates of para-Cl BODIHY.

29
Coordinates from ORCA-job pCl

F	-1.08080851408104	0.91229696629693	-1.16497966834548
F	-1.53185671273358	-1.19887533148372	-1.98902592919119
B	-1.10584290319460	-0.45644282113870	-0.89182257160851
N	-1.98007278859908	-0.74235569955218	0.38364655209850
N	-1.49531515971821	-0.86402169779917	1.59553408070623
N	0.38151123756865	-0.91238040441900	-0.54122269680042
N	0.64918595161395	-1.10035411157551	4.24196446943466
C	-0.17862399696762	-0.93602605617860	1.78890382649021
C	0.81562655056560	-1.02646815334847	0.74972291798279
C	2.17791846221187	-1.30113465332783	1.00072225814113
C	3.05383563230769	-1.46213262529000	-0.05661026873532
C	2.57755229635619	-1.36108702759221	-1.37702277622399
C	1.24086999061599	-1.08560483974441	-1.57834498113977
C	0.24787641112681	-1.02512495495766	3.14751113216190
C	-3.40339479547531	-0.79025442378892	0.29822079274486
C	-4.06060250827433	-0.21841590853129	-0.80321457051814
C	-5.45087105523857	-0.26090276363429	-0.88616453561092
C	-6.18410615240857	-0.87543341003870	0.13007477274485
C	-5.54255390964179	-1.44570191152256	1.23269798112271
C	-4.15375375856045	-1.40710927659887	1.31240462464018
Cl	-7.93736361994884	-0.93402001410781	0.02189322424303
H	2.50988366938917	-1.39080271245026	2.03514729218739
H	4.10695236654848	-1.67379241034345	0.13492403330824
H	3.23728338234664	-1.48858902806575	-2.23384968643609
H	0.79561907900078	-0.98633495670525	-2.56752439793736
H	-3.49242132590745	0.27818908650622	-1.58802799574993
H	-3.63869945766542	-1.84824958384399	2.16485600183229
H	-5.96362771749342	0.18717472131741	-1.73710218589029
H	-6.12701981874357	-1.92652549408196	2.01720772434842

Table B.14. PBE/TZVP-optimized Cartesian coordinates of dimeta-Cl BODIHY.

29
Coordinates from ORCA-job m2Cl

F	-1.03883502803491	0.96521241276602	-1.05901799437314
F	-1.58459156386207	-1.11496051816359	-1.90468934605684
B	-1.08982944047324	-0.40709618026198	-0.81572574815571
N	-1.91236890180552	-0.69889981324700	0.49307315647482
N	-1.37703604478911	-0.85308424220930	1.68118800346482
N	0.40040151780874	-0.90833682385065	-0.53960276073693
N	0.88114186554574	-1.17779905491567	4.22218836967068
C	-0.05756405022688	-0.95628745520318	1.81384071878970
C	0.88846923035740	-1.05225938580357	0.72773459243450
C	2.25230018844050	-1.36289841382447	0.91461025096090
C	3.07679990300135	-1.52773279923701	-0.18380037848097
C	2.54516769096635	-1.39525525267127	-1.47898041482558
C	1.20733201806049	-1.08526988665459	-1.61682566742533
C	0.42886789833650	-1.07649269644837	3.15021959892589
C	-3.33950472881645	-0.71128942590202	0.47452500404425
C	-4.02191921959248	-0.09770200210689	-0.58728796671451
C	-5.41452391218943	-0.12363496220812	-0.58495486730729
C	-6.14811041830562	-0.73531641494898	0.43225066135953
C	-5.43487858303616	-1.33205787725697	1.47433951124309
C	-4.04487773366375	-1.33765126119598	1.51370195574028
Cl	-6.32890887288004	-2.11555933309289	2.76913449188111
Cl	-6.28075534829020	0.64995238522726	-1.90367438022913
H	2.62750877801680	-1.47672115752446	1.93175441619116
H	4.13199220142430	-1.76664950957754	-0.04278457685786
H	3.16300676223761	-1.52475224834509	-2.36620029583546
H	0.72099984018415	-0.95991713731734	-2.58347050283804
H	-3.48518885420968	0.40703890334523	-1.38700916493872
H	-3.50648849944786	-1.80878185503285	2.33295909769033
H	-7.23659204275650	-0.74607298333875	0.41463758190443

Table B.15. PBE/TZVP-optimized Cartesian coordinates of meta-Cl BODIHY.

29
Coordinates from ORCA-job mCl

F	-1.02667037721651	0.98113391663904	-1.05938480951745
F	-1.58530419481326	-1.09108220220200	-1.91746535988428
B	-1.08950901700233	-0.39257790694510	-0.82227384267523
N	-1.91720783698336	-0.68332620713396	0.48317462271417
N	-1.38463441929754	-0.82806297571884	1.67256366899412
N	0.39557960973191	-0.90623280386839	-0.54391009106832
N	0.87141710186719	-1.12469588540781	4.22058304719693
C	-0.06369750628461	-0.92874594198261	1.80876992394832
C	0.88217847333215	-1.03961488152073	0.72583851223311
C	2.24465193315833	-1.35693290917154	0.91609631397376
C	3.06790841334664	-1.53998778919358	-0.17975736954260
C	2.53706886156975	-1.41983671747856	-1.47707184915992
C	1.20176055580899	-1.10224304456162	-1.61863442196042
C	0.42061898061244	-1.03488002011112	3.14682517794769
C	-3.34496012336534	-0.70250833290820	0.45669670120870
C	-4.03534575634344	-0.08709013176605	-0.59929181992932
C	-5.42948839819999	-0.11194525504127	-0.60940095933212
C	-6.14987940947296	-0.73867226931953	0.40977825730092
C	-5.44081170589159	-1.34258417792481	1.45095796280822
C	-4.05063341740403	-1.33996606446628	1.48956390568148
Cl	-6.32799891357141	-2.14327967228147	2.74401546164475
H	2.61880046469442	-1.46173652063567	1.93461411838607
H	4.12152095175080	-1.78406666820186	-0.03575249174958
H	3.15420147989712	-1.56426319345134	-2.36246476319641
H	0.71638982342753	-0.98475048043086	-2.58672359701077
H	-3.49041330264820	0.42194516819072	-1.39198148661082
H	-3.51032009245411	-1.81479209140707	2.30614271016718
H	-7.23969535737972	-0.75986235282962	0.40198432976776
H	-5.96745204986887	0.37170008113015	-1.42639848133595

Table B.16. PBE/TZVP-optimized Cartesian coordinates of para-Br BODIHY.

29
Coordinates from ORCA-job pBr

F	-1.03529789719348	0.94833099354549	-1.08901449414623
F	-1.55959555001194	-1.13736130239639	-1.93486358180636
B	-1.07778646922340	-0.42470064961440	-0.84157234888346
N	-1.91020972971949	-0.71788973391932	0.46034161657753
N	-1.38283363306894	-0.86035847235660	1.65246695370088
N	0.41175127027882	-0.91534065811233	-0.55412618169481
N	0.85748042869161	-1.15390033231489	4.21305511817986
C	-0.06128546339661	-0.95292598288570	1.79528152005682
C	0.89203354610187	-1.04893012934782	0.71853234674847
C	2.25757364205239	-1.34934328776009	0.91576569594182
C	3.09008543273284	-1.51595981180644	-0.17545029130903
C	2.56604855787267	-1.39567630230869	-1.47587042024128
C	1.22786341361373	-1.09493453314551	-1.62451384643593
C	0.41492542917992	-1.06190944537437	3.13593821717712
C	-3.33584643762376	-0.74626514299992	0.42717329729604
C	-4.02538664490028	-0.14767538202936	-0.63953141447576
C	-5.41887437859119	-0.17096178524389	-0.67141481620053
C	-6.12158495199820	-0.79273571123138	0.36174725789339
C	-5.44786270482704	-1.39025301461413	1.43023210265471
C	-4.05618328700392	-1.37033114822193	1.45848392527725
Br	-8.03479867463753	-0.82930254178503	0.31457903505550
H	2.62658749993389	-1.45427414934731	1.93612034472115
H	4.14588481220552	-1.74707549665686	-0.02597867668339
H	3.19078760738714	-1.52739030002271	-2.35790272668946
H	0.74755425360236	-0.97896284572158	-2.59529940870970
H	-3.48088220815463	0.35475471017050	-1.43743577078194
H	-3.51587162202173	-1.83274346176428	2.28367986466389
H	-5.95278084351573	0.29944202795801	-1.49682172704901
H	-6.00520652676486	-1.87768417769307	2.23012974416248

Table B.17. PBE/TZVP-optimized Cartesian coordinates of dimeta-Br BODIHY.

29
Coordinates from ORCA-job m2Br

F	-1.03853550449361	0.95898698320437	-1.06189083452625
F	-1.57751874384470	-1.12442301051357	-1.90377185956949
B	-1.08516128538758	-0.41295570544975	-0.81590447217486
N	-1.90760157311335	-0.70495974467024	0.49306295355080
N	-1.37194046699502	-0.86401432506030	1.68052840266789
N	0.40673669214063	-0.90881843524485	-0.53977860492227
N	0.88643281553690	-1.20452483707841	4.21986260248843
C	-0.05290747263660	-0.96970980449760	1.81299163940419
C	0.89413491048892	-1.05875342409513	0.72696382975356
C	2.25859484122392	-1.36662923091009	0.91354593573615
C	3.08480711492860	-1.52174078605728	-0.18506193443719
C	2.55415562879307	-1.38215139254564	-1.47986300431018
C	1.21542154726246	-1.07586013333267	-1.61724855773773
C	0.43307585441797	-1.09752704250409	3.14889976019766
C	-3.33489963671745	-0.71251717899432	0.47547006716876
C	-4.01663043269775	-0.10493613297256	-0.59097780478697
C	-5.40911486030112	-0.12741025729510	-0.58787652912898
C	-6.14591242643231	-0.72670603862525	0.43476152414748
C	-5.43268971573580	-1.31652878089730	1.48138227665980
C	-4.04307745301319	-1.32785065331248	1.52027407708924
H	2.63288927521201	-1.48565911343725	1.93040719440985
H	4.14061764228541	-1.75828683992783	-0.04465000070137
H	3.17346295395513	-1.50371975483957	-2.36719177163450
H	0.72983353281792	-0.94602168285365	-2.58366561336857
H	-3.47558907250428	0.38982061970627	-1.39386901411950
H	-3.50209267766500	-1.79355803484486	2.34084226092606
H	-7.23429744017949	-0.73412586795223	0.41749476265232
Br	-6.40957554076231	-2.15149789382706	2.90175842048385
Br	-6.35060384958338	0.70180351182845	-2.03436236391813

Table B.18. PBE/TZVP-optimized Cartesian coordinates of meta-Br BODIHY.

29
Coordinates from ORCA-job mBr

F	-1.04917164770672	0.92992863387484	-1.09884123569828
F	-1.57799594293258	-1.16866356082547	-1.90991831114226
B	-1.09129302649750	-0.43900791306167	-0.83079375891873
N	-1.91383045352842	-0.71341179314953	0.48136442799905
N	-1.37848117894798	-0.83550098908923	1.67209937610710
N	0.40148441484590	-0.92447390188276	-0.54447578536245
N	0.88348420275778	-1.07542673023661	4.22137891411270
C	-0.05675335975844	-0.92149723110347	1.80846317307151
C	0.89017019001052	-1.03508246235935	0.72655366449342
C	2.25686721090042	-1.33104300099629	0.92047595215310
C	3.08252906450053	-1.51599277080184	-0.17331590333885
C	2.54998475271463	-1.41882675808992	-1.47182440181482
C	1.21033124542735	-1.12194551518703	-1.61692555587396
C	0.42974425884273	-1.00497717266196	3.14741776804556
C	-3.34141604195123	-0.74472994392300	0.45706281364492
C	-4.03766435238790	-0.15526370840589	-0.60961127422532
C	-5.43147957325246	-0.19179408199441	-0.61869185259874
C	-6.14687396754108	-0.80422259192536	0.41325311514966
C	-5.43151152142424	-1.38233805044746	1.46459515685851
C	-4.04167007909116	-1.36878849041929	1.50275713428799
H	2.63244703426102	-1.41783801047259	1.94014962344058
H	4.13950001786159	-1.74358397956866	-0.02669636702040
H	3.16908039487278	-1.56532936879061	-2.35550679029925
H	0.72328441877614	-1.02353649999966	-2.58633058313328
H	-3.49780125240944	0.34308496753799	-1.41256163477017
H	-3.49306156249627	-1.82326093486667	2.32523308493541
H	-5.97378963433954	0.27198778407684	-1.44441381767195
H	-7.23641892405486	-0.83161313094135	0.40297815923129
Br	-6.39118635345159	-2.23265568928958	2.89144654933764

Table B.19. PBE/TZVP-optimized Cartesian coordinates of para-OH BODIHY.

30
Coordinates from ORCA-job pOH

F	-1.06902496740342	0.94836479452051	-1.14179820758846
F	-1.55258853380546	-1.14784874556099	-1.98447415728942
B	-1.11109876168886	-0.42354170071460	-0.87987156498598
N	-1.98816968372544	-0.70571447783500	0.39531924070621
N	-1.50266892406600	-0.83883928854865	1.60397176322496
N	0.36720843910059	-0.90513599418934	-0.53843248459373
N	0.64724872924650	-1.11839511114680	4.24315370147987
C	-0.18317566615295	-0.93095509721201	1.79293573902368
C	0.80388945402705	-1.03136322695390	0.75217764194094
C	2.16453056406824	-1.32725751711765	0.99570017402049
C	3.03326844002114	-1.49790612125222	-0.06467284639276
C	2.55326905051546	-1.38577980792673	-1.38428944874953
C	1.22107221935300	-1.08919818714478	-1.57924912532553
C	0.24494353774668	-1.03282240601078	3.14951896640964
C	-3.41088889202833	-0.74090024455910	0.30748012347550
C	-4.06423371960229	-0.15290225144233	-0.78689776178927
C	-5.45403955273282	-0.18201843210943	-0.86758926894412
C	-6.20939575415684	-0.79901990640784	0.13897690892437
C	-5.55804393836178	-1.38512547805675	1.23468659110781
C	-4.17232979643502	-1.36138653204001	1.31365884071331
O	-7.57703447411016	-0.86273667464960	0.11002101658125
H	2.49875058858404	-1.42575111363528	2.02861748712848
H	4.08381525710844	-1.72601757280897	0.12202240964568
H	3.20789133249749	-1.52064449203738	-2.24386949123230
H	0.77386037417022	-0.97946019812339	-2.56639642299574
H	-3.49192900158221	0.34717425117482	-1.56648265576838
H	-3.66098075732996	-1.81669307045847	2.16104099883340
H	-5.95301024933254	0.28368074467482	-1.72225847580314
H	-6.15426874000195	-1.86886942193685	2.00931527648036
H	-7.89578092692282	-0.42595119549129	-0.69964986323758

Table B.20. PBE/TZVP-optimized Cartesian coordinates of BODIHY.

29
Coordinates from ORCA-job BF2

F	-1.07910501641908	0.92823998998185	-1.15472807166406
F	-1.54184833488529	-1.17531235700136	-1.99451658851993
B	-1.11553139466417	-0.44237326890705	-0.89159248802755
N	-1.99182775040532	-0.73087496700014	0.38140213964495
N	-1.50856938461816	-0.84070180534578	1.59383850767861
N	0.36925024803669	-0.90945214077731	-0.54006765374669
N	0.63650194819589	-1.04027155840878	4.24436004973648
C	-0.19094056834157	-0.90507281731252	1.78946436679168
C	0.80317556907380	-1.01012088470968	0.75212914944713
C	2.16436553024692	-1.28931558810398	1.00578733160868
C	3.03844428723910	-1.46907127906253	-0.04984273115283
C	2.56174228700228	-1.38295656960767	-1.37133898103253
C	1.22663345890671	-1.10195761255438	-1.57526244613800
C	0.23485102728467	-0.97797220593304	3.14913876788554
C	-3.41716518967610	-0.79394928745553	0.29258566588356
C	-4.07603028486110	-0.20135816888692	-0.79657113047148
C	-5.46734338429054	-0.26405078780287	-0.87560639356612
C	-6.20906306686718	-0.90968056406157	0.11737332944095
C	-5.54651270171954	-1.49631784400486	1.20115118288298
C	-4.15728632604346	-1.44633531895834	1.29176108598715
H	2.49631083004307	-1.36785790603918	2.04113286378048
H	4.09050884958426	-1.68416774567605	0.14374061270915
H	3.22016657092505	-1.52565825923599	-2.22677737364000
H	0.78089763657913	-1.01250573840807	-2.56517012553676
H	-3.50726704316992	0.32019554140646	-1.56487730857150
H	-3.63221740552485	-1.90238987575952	2.13034724138227
H	-5.97422171123830	0.20327053404599	-1.72180723438776
H	-6.11573688628408	-2.00721652133898	1.98013598534719
H	-7.29751324510890	-0.95706548808220	0.04732827124839

Table B.21. PBE/TZVP-optimized Cartesian coordinates of BCl_2-hydrazone.

29
Coordinates from ORCA-job BCl2

Cl	-0.96346050562657	1.35749295685579	-1.30658754255906
Cl	-1.70070729192461	-1.45049561354547	-2.37376057496634
B	-1.08925327139566	-0.47538726533496	-0.91607655001505
N	-1.98505325879795	-0.69924799661366	0.33494722800753
N	-1.51659852802193	-0.75270388441024	1.55941549633544
N	0.37602673503967	-0.95544829225421	-0.55191408964835
N	0.59707156258346	-0.82910164063722	4.24093676217133
C	-0.20477978038099	-0.81118805034825	1.77387251411760
C	0.79411900045866	-0.98939334635735	0.75238462874051
C	2.14479050684060	-1.27302703918730	1.04479995830958
C	3.03937878260080	-1.51482194282711	0.01955850984008
C	2.58621322902281	-1.48573008100271	-1.31101625307150
C	1.25921796056815	-1.20325011627910	-1.55673604256875
C	0.20718469653115	-0.81718824655114	3.14001353800128
C	-3.41754578223541	-0.78021140835890	0.26368684017106
C	-4.12904484362576	-0.06142016064857	-0.70695311984488
C	-5.52087903797406	-0.15241770510738	-0.74349945672006
C	-6.20718014461630	-0.95067639107426	0.17587978448874
C	-5.49081499777974	-1.66090574335338	1.14440740508325
C	-4.10000451972921	-1.58165254335673	1.19041694266248
H	2.45347280002648	-1.30169985100749	2.08984493201434
H	4.08498349201887	-1.73015776702456	0.24448663723721
H	3.25726732783100	-1.67123949127377	-2.14821752074212
H	0.84024163324658	-1.15266668187132	-2.56010053279870
H	-3.60223358586765	0.58070049915719	-1.41109318126516
H	-3.53258019921078	-2.13514748530542	1.93813253927554
H	-6.07199279806095	0.41412529683783	-1.49617064689139
H	-6.01689235966959	-2.29035164962929	1.86458296754755
H	-7.29627827685109	-1.01908886049099	0.13827685108782

Table B.22. PBE/TZVP-optimized Cartesian coordinates of meta-NO_2 BODIHY.

31
Coordinates from ORCA-job mRNO2

F	-0.35252884310299	-2.18618048083917	-0.58827378138022
F	0.08147295824252	-1.67210294778435	1.62290431955647
B	0.10235848379241	-1.22001197241535	0.30283335901653
N	-0.71004857029827	0.11838626413727	0.17113646026601
N	-0.17777303709303	1.31571701947190	0.26006798449116
N	1.61337171095431	-0.87065038904804	-0.06934951088403
N	2.08940183860987	3.85778189218138	0.54545118040514
C	1.14421572056601	1.45766110866661	0.27004183856623
C	2.10106019561707	0.40067696922728	0.03945865444249
C	3.47837005336945	0.63428167754169	-0.15835839601110
C	4.31705074871965	-0.42091467898536	-0.47034518666643
C	3.78641532822659	-1.71696905847872	-0.59914447379455
C	2.43464131687540	-1.90245811618338	-0.39139592792193
C	1.63173315833428	2.79044189319574	0.42348680676505
C	-2.13046202703870	0.10425122808130	0.04709440957612
C	-2.86205734513090	-1.02345935385942	0.45751259582509
C	-4.25174223382141	-1.03458639317102	0.33094219924474
C	-4.93241337736734	0.06221876906697	-0.19818418088865
C	-4.18103510880590	1.16862816965897	-0.59493119287128
C	-2.79591946085949	1.21472885018373	-0.48932332979960
N	-4.88772031618326	2.34791790858866	-1.17526609559254
O	-4.19944906464250	3.30589061298367	-1.53717481647183
O	-6.11857596503213	2.28769361635731	-1.25627741329164
H	3.85273713470059	1.65415845745327	-0.06944934718062
H	5.38288730539645	-0.24373931803830	-0.62279159996106
H	4.41575988968429	-2.56930312354641	-0.85040824392528
H	1.94765284110395	-2.87415617085218	-0.46613597473135
H	-2.34813768087032	-1.88114720596353	0.88893457148241
H	-4.81108604943020	-1.91218294221574	0.65772551739130
H	-6.01568706196769	0.07673255933576	-0.30914704284845
H	-2.24760625254872	2.09607392424947	-0.81589006780816

Table B.23. PBE/TZVP-optimized Cartesian coordinates of dimeta-NO_2 BODIHY.

33
Coordinates from ORCA-job m2NO2

F	-0.35379773299243	-2.18334204718464	-0.53709312977355
F	0.06233855621945	-1.61401125465788	1.66392842595303
B	0.09927762768946	-1.19999177759765	0.33457190787778
N	-0.70727441838407	0.13795606899015	0.15548077998516
N	-0.16963094478504	1.33942695073142	0.16476835117162
N	1.61704129748499	-0.87031810716004	-0.03325622859425
N	2.09979804600773	3.88987743131786	0.24652419199230
C	1.14899763449213	1.47792200454229	0.14598557773730
C	2.10601308068710	0.40345055415666	-0.01221389852153
C	3.48323419393687	0.62249543483199	-0.21661041453372
C	4.32396005608267	-0.45409055879043	-0.44322164160736
C	3.79349416678630	-1.75468766205674	-0.47788502622283
C	2.43901509307625	-1.92400607631155	-0.26846713949082
C	1.64038684964755	2.81748374969200	0.20272561698734
C	-2.12696102225384	0.12694885730617	0.07442150162634
C	-2.84355320384668	-1.00577296388712	0.49170661542182
C	-4.23158223538205	-0.99018726147807	0.39280621182559
C	-4.94904608635873	0.09730007199982	-0.09463873707759
C	-4.20203070202326	1.20247815000757	-0.49191547671491
C	-2.81443124564504	1.24396355772838	-0.42621705690914
N	-4.92620445303723	2.39473503213714	-1.03531809024213
N	-4.98949011366315	-2.20004609197360	0.84236177897607
O	-4.24271559849528	3.35373128623463	-1.39895073848995
O	-6.15621368686600	2.33289251805388	-1.08089317013930
O	-4.33536289733059	-3.13910961992844	1.29993266267655
O	-6.21572213853304	-2.17224257820762	0.72199221152352
H	3.85866831608618	1.64572585496169	-0.19907467343879
H	5.39106041485433	-0.28902314068549	-0.60020451455936
H	4.42377328808233	-2.62392577687680	-0.65919640622338
H	1.95213127125404	-2.89872316218201	-0.27513517763337
H	-2.35068642816201	-1.88578819691975	0.90065645744008
H	-6.03502419700108	0.08382124394334	-0.16397393664021
H	-2.27669425262787	2.13058342826286	-0.75632359938231

Table B.24. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 BODIHY.

31
Coordinates from ORCA-job pNO2

F	-1.60470739986312	-3.05345354956410	-0.37962368892859
F	-1.22204222478440	-2.53055036550212	1.83918421622555
B	-1.15852457914806	-2.09001527945549	0.51799191456650
N	-1.94282162223507	-0.73561724681343	0.35348419042993
N	-1.38174352570080	0.45253479256839	0.40452927297749
N	0.37172391787765	-1.77864824248821	0.18357082466051
N	0.92019648516080	2.96467341026790	0.60055290046140
C	-0.05875374034822	0.56854268233028	0.41988626578941
C	0.88266958780991	-0.51479988805114	0.25152822360940
C	2.26840852461176	-0.31563167538004	0.08069649146504
C	3.09366891357978	-1.39985247346510	-0.16176770388771
C	2.54057460523755	-2.68950295628653	-0.24629187031655
C	1.17979484419523	-2.83950547211607	-0.06818610760165
C	0.45031805819230	1.89858475441958	0.52134213047359
C	-3.36077354558704	-0.71988765169662	0.23444646894542
C	-4.10702909587656	-1.85270426894697	0.60826134486125
C	-5.49224587826980	-1.84017317088649	0.48784086329092
C	-6.12527040219642	-0.69790753454689	-0.00334963530533
C	-5.40043280753437	0.43686059328245	-0.37424838707200
C	-4.01670087971123	0.42328177207225	-0.25930594795430
N	-7.60117531537472	-0.68989127369569	-0.13685350306558
O	-8.21082195854941	-1.71819922946056	0.17900864858605
H	2.66132993946982	0.69968636587264	0.13604793374387
H	4.16657217300707	-1.25004340591171	-0.29299791241037
H	3.15874837548129	-3.56429941060782	-0.44190005875870
H	0.67531328024156	-3.80408445467320	-0.11405102443151
H	-3.60997726126553	-2.73418102145625	1.00878347840024
H	-3.42814274540677	1.29296767978080	-0.54674871561236
H	-6.09084649372041	-2.70329687846167	0.77597951940178
H	-5.92837291106589	1.30893940096815	-0.75846115559282
O	-8.13249677922694	0.34510254090366	-0.55645153795085

Table B.25. PBE/TZVP-optimized Cartesian coordinates of para-CO_2H BODIHY.

32
Coordinates from ORCA-job pCO2H

F	-0.77712327358763	1.14203924342672	-0.94548589072276
F	-1.48319620265529	-0.80084873952257	-1.98056789503048
B	-0.97267917780488	-0.23336078828482	-0.81860191083490
N	-1.87430039048387	-0.54121930349876	0.43334916922868
N	-1.39721155370894	-0.80448974144238	1.62849034924531
N	0.44924633527244	-0.90049626663676	-0.53302466056915
N	0.73335878588629	-1.45590143769762	4.21778031321414
C	-0.09727381803071	-1.03044563462046	1.79698908276337
C	0.87607834607413	-1.15713185586110	0.73891001939214
C	2.19696982684405	-1.60643097866743	0.95024561300283
C	3.04091325178982	-1.79259304349509	-0.12982510489498
C	2.57014528729121	-1.54308733145897	-1.43145383203339
C	1.27347448272555	-1.09872096925319	-1.59304436441600
C	0.32865758699598	-1.26004817446519	3.13987145784393
C	-3.29301902416695	-0.42470023070023	0.36879772019553
C	-3.88896290966909	0.29359383659674	-0.68219852563433
C	-5.27429111935248	0.40796903291969	-0.74215484141754
C	-6.08452601014455	-0.18793521782025	0.23722876882119
C	-5.47962644721473	-0.90151405784960	1.28570592621430
C	-4.09899347404480	-1.02472483161488	1.35317966526220
C	-7.56645009067980	-0.09570097622592	0.21235239312386
O	-8.31323150344361	-0.58416216719757	1.04425540784005
H	2.52331698025997	-1.80883168161530	1.97042109812175
H	4.06320015878030	-2.13839769202014	0.02948411187632
H	3.20428343004360	-1.68595043623149	-2.30508845427501
H	0.83645720963714	-0.87660136008826	-2.56567512692496
H	-3.27336970813455	0.77679137985397	-1.43873341233708
H	-3.62668651009594	-1.57882377780716	2.16331528441707
H	-5.73487717028901	0.97035858894355	-1.55435790265109
H	-6.11667576133664	-1.36505593343923	2.04029884151299
O	-8.03632138560091	0.61150260720845	-0.86183135296249
H	-9.01077888115609	0.60956705356526	-0.76877525437149

Table B.26. PBE/TZVP-optimized Cartesian coordinates of para-CO_2CH_3 BODIHY.

35
Coordinates from ORCA-job pCO2CH3

F	-1.00763039670735	1.15292366110491	-0.85345017187299
F	-1.52015553554657	-0.80300199403366	-1.97336999074551
B	-1.11122792834807	-0.23650554102542	-0.77100481870483
N	-2.05499072597948	-0.64073933626521	0.42024465665084
N	-1.62751487374277	-0.91723452467421	1.63051723646295
N	0.33587567362837	-0.81264838113085	-0.42507910427496
N	0.39557878536259	-1.52590659693802	4.31384455079305
C	-0.32639082563168	-1.06732288693375	1.86431287928603
C	0.70885108024406	-1.09058873594617	0.85953039666748
C	2.04347241624254	-1.46058097068068	1.12980408168026
C	2.95601292303330	-1.54790733449490	0.09413913467633
C	2.54196142262165	-1.27793955656947	-1.22258477895324
C	1.22898901860251	-0.91407461532914	-1.44236621965942
C	0.03974208112467	-1.31937323012821	3.22071697986934
C	-3.47334416831733	-0.61440626379552	0.27928219222928
C	-4.05958693236122	0.08972186997103	-0.78510255692015
C	-5.44636887058483	0.13245445722790	-0.90483344787704
C	-6.27050439593265	-0.53934369432395	0.01066745049589
C	-5.67133174745894	-1.26859986413434	1.05150405492294
C	-4.29048071473411	-1.29596668143863	1.19799897634708
C	-7.75767906042044	-0.62884252369854	-0.15598578249340
O	-8.35491846308566	-1.68210317600402	-0.06288892842061
O	-8.46823222373314	0.50319475295521	-0.44986879685582
C	-7.93095639453654	1.82114895707219	-0.21441724660380
H	2.32601005064511	-1.68220338105524	2.15896886772271
H	3.98896723615815	-1.83266581779672	0.29954682493106
H	3.23129977879572	-1.34403336139932	-2.06302084837430
H	0.83202596518143	-0.68359338453133	-2.43009164440405
H	-3.43898563587241	0.61618431033806	-1.50789688581978
H	-3.82785913035955	-1.84879612424349	2.01469316752967
H	-5.88558357789475	0.67823854990635	-1.74145395859584
H	-6.30581248784110	-1.82770089816064	1.74086191888712
H	-8.77300161500745	2.43002462903039	0.13773878695556
H	-7.54147981065455	2.25061083343314	-1.14854470806474
H	-7.13688926488949	1.81155844669222	0.54451996153289

Table B.27. PBE/TZVP-optimized Cartesian coordinates of AlF_2 hydrazone.

29
Coordinates from ORCA-job AlF2

F	0.48152712237284	2.95920966572783	-0.47507141549438
F	0.20877981241736	1.36267586560918	-2.81989277270338
Al	0.34947881763294	1.40933436022716	-1.13151976323432
N	-0.82117066453457	0.28714888271908	-0.12152752289222
N	-0.42599148124366	-0.56514886185292	0.79615432708774
N	1.97125429887895	0.43407537478762	-0.72071327633261
N	1.20478903099985	-2.57893184898489	2.94742117764762
C	0.85022133861752	-0.88693657998446	1.00440094926863
C	2.05198630749487	-0.49209852010391	0.28636039185505
C	3.30473717272146	-1.07839976322933	0.58494884771030
C	4.42852017627236	-0.72861124237724	-0.14163724590129
C	4.31904857389472	0.20507096151991	-1.18620534933811
C	3.07949876358162	0.75760870144392	-1.43911528823451
C	1.02460524986700	-1.82125435394360	2.07681231290119
C	-2.23389125669437	0.47213029151213	-0.19771357542773
C	-2.73851811293649	1.21250078154834	-1.27947275158472
C	-4.10993259976566	1.44710062112600	-1.38492864025441
C	-4.98726316668517	0.94358922858803	-0.42167515121159
C	-4.48034177879013	0.19848863470782	0.65052366031796
C	-3.11341618873653	-0.03952965281981	0.77056297829126
H	3.36253370665760	-1.80879370429470	1.39191592015458
H	5.39200455018422	-1.18270708516251	0.09454869441027
H	5.17884223635508	0.49880127739569	-1.78645615665057
H	2.92541933610410	1.48543902417927	-2.23830449692323
H	-2.07196234437424	1.57995383790286	-2.06608486245380
H	-4.49116506574647	2.02117781951618	-2.23124053076813
H	-6.05944984680539	1.12977359380266	-0.50422333951010
H	-2.71069105532218	-0.61127439236426	1.60575867637639
H	-5.15886093341762	-0.19742391619607	1.40855820289409

Table B.28. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 AlF_2 hydrazone.

31
Coordinates from ORCA-job AlF2_pNO2

F	0.33661041919763	2.10876874494086	-2.10170978057597
F	-0.01024255040828	-0.60823742475105	-2.90853326950410
Al	0.24797162459948	0.48085952295735	-1.64709042946344
N	-0.82932713604930	0.06900505078701	-0.12207218657134
N	-0.42247269626023	-0.67225986813236	0.89107329919099
N	1.94597314948414	0.06753280804364	-0.81519193907077
N	1.22179605297659	-2.51660935387597	3.17561404498358
C	0.84312282227439	-1.02806015553032	1.07600567073326
C	2.04888341630832	-0.66249264498619	0.33892064517419
C	3.32856851636803	-1.01198302776370	0.82421911370948
C	4.46460367051860	-0.60730646797765	0.14356500052059
C	4.33834814268526	0.16075689208101	-1.02420186382682
C	3.06714205014507	0.47395479190949	-1.46638109252444
C	1.03228320028584	-1.84733717643873	2.23772457549855
C	-2.21588066677112	0.37415433366428	-0.10717539915390
C	-2.68673850096362	1.33075091912588	-1.02666110337569
C	-4.03843349090378	1.65430855824847	-1.07083365900824
C	-4.91520464232661	1.01993693548112	-0.19090930600410
C	-4.46532672925640	0.07236460924527	0.73282790823395
C	-3.11576724062990	-0.25192290740597	0.77572522723362
N	-6.35544434051356	1.36340448189935	-0.23337390962317
H	3.40143357670240	-1.59498056162070	1.74190656796847
H	5.45106841189760	-0.88084238447871	0.52063732491932
H	5.20815881162497	0.50528436839721	-1.58119998453026
H	2.89978138311161	1.07069255602243	-2.36522084875325
H	-1.99321213755245	1.85843928154086	-1.68794240413511
H	-4.42117052580603	2.39632503597540	-1.77088338716447
H	-2.74197109701304	-0.99078598867482	1.48304723150647
H	-5.18226123572865	-0.40321878098704	1.40120351846369
O	-7.10439625479282	0.78682515226739	0.56426928156639
O	-6.71940930220417	2.20524003603618	-1.06269181341748

Table B.29. PBE/TZVP-optimized Cartesian coordinates of $AlCl_2$ hydrazone.

29
Coordinates from ORCA-job AlCl2

Cl	0.39243669859509	3.27599441476155	-1.06750206810668
Cl	0.20007554126456	0.60641207619012	-3.46939600653155
Al	0.32456045811626	1.16544959612644	-1.40569986804393
N	-0.83779193094986	0.19031240167388	-0.23261175342624
N	-0.45019677050022	-0.73455027716865	0.61568574877800
N	1.98626380556467	0.41833065871541	-0.72399748689940
N	1.11676964287174	-2.98777665740125	2.56044092758133
C	0.81738965289302	-1.09202497263909	0.80457546201591
C	2.04319827855008	-0.57702887828775	0.21627379764143
C	3.30453166035990	-1.07696840458483	0.61695496378320
C	4.46539276556721	-0.56369573430213	0.06823582414716
C	4.38569663403173	0.46174913480387	-0.88847415120086
C	3.13540105444820	0.91884601390459	-1.25213147837182
C	0.96323652492639	-2.13752693940048	1.77441861320143
C	-2.24299173850001	0.44941353264129	-0.21598602628588
C	-2.78409107478825	1.24778273384246	-1.23567331972272
C	-4.14688272083499	1.54547859353506	-1.24382031708771
C	-4.98224704965001	1.04954013008942	-0.24031892997186
C	-4.44076787981425	0.24920479062143	0.77299960490172
C	-3.08094383489947	-0.05092341096077	0.79441169838339
H	3.33929239250344	-1.86899652451072	1.36493714461875
H	5.43470665173212	-0.95427865330010	0.38113489564939
H	5.27541242918587	0.89455887243769	-1.34276699244774
H	3.00751774433635	1.71351136877340	-1.98980936241002
H	-2.15620029321074	1.62669932239198	-2.04719668134836
H	-4.55406966014416	2.16554864517639	-2.04428521008465
H	-6.04747728335365	1.28629390242122	-0.24553073884416
H	-2.65286182457272	-0.66716405446270	1.58399332966482
H	-5.08476787372831	-0.14122268008775	1.56332237841705

Table B.30. PBE/TZVP-optimized Cartesian coordinates of BBr_2 hydrazone.

31
Coordinates from ORCA-job AlCl2_pNO2

Cl	0.44336509862424	3.28131710488678	-1.03093753659095
Cl	0.19850667108394	0.63345118491183	-3.45802718825338
Al	0.35824635282736	1.17991332896524	-1.39777681508934
N	-0.80811462502772	0.21284031654424	-0.21949007448105
N	-0.41681052328368	-0.69900407501994	0.64983970427629
N	2.01877740608490	0.40543419376469	-0.73575851575985
N	1.14794534737547	-2.91048951533139	2.63873420547943
C	0.84546220476337	-1.05505802223387	0.83987742256939
C	2.07525584548337	-0.56717042086937	0.22533267954422
C	3.33104170679762	-1.07815668689784	0.62109268782830
C	4.49324300720473	-0.59657306742197	0.04281202333106
C	4.41504840660139	0.40365062622068	-0.93761224780866
C	3.16589975219470	0.87294249688589	-1.29477874939751
C	0.99309051916688	-2.07843820935252	1.83440145474181
C	-2.20953891631101	0.45523377724413	-0.20649397761778
C	-2.75242969734092	1.28629836518199	-1.20345434168622
C	-4.11343079935627	1.56886889912937	-1.22197922182398
C	-4.93184709913230	1.01436333537520	-0.23812826607195
C	-4.41407761738050	0.18351939892858	0.75893806964879
C	-3.05381343622350	-0.09486983832924	0.77688213597889
N	-6.38214001019468	1.31617414213298	-0.24976423687630
H	3.36473286147970	-1.85257062086427	1.38735247829118
H	5.46060548582556	-0.99473663853923	0.35196095645327
H	5.30444621630906	0.81039264311838	-1.41610814514323
H	3.03969917155473	1.64975771071411	-2.05156868502123
H	-2.12214304458790	1.70773904430404	-1.99083927290145
H	-4.54888689097576	2.20855654308378	-1.98879176623598
H	-2.62795713858505	-0.73577918123911	1.54698802367156
H	-5.08439328072344	-0.22904078527582	1.51213947223705
O	-7.07864409506774	0.80821154844219	0.63681554186410
O	-6.80590647718654	2.05645740654046	-1.14482589215651

Table B.31. PBE/TZVP-optimized Cartesian coordinates of BBr_2 hydrazone.

29
Coordinates from ORCA-job BBr2

Br	0.76248800375744	3.00495454717129	-0.45503019861282
Br	-0.16872775482886	0.86792348895124	-2.90366982850467
B	0.39077410241228	1.04443946837432	-0.95940855814224
N	-0.66910910245678	0.43526512386068	-0.01241175121802
N	-0.36783772878183	-0.22276556865786	1.08412685252327
N	1.75434716615450	0.28202877730539	-0.73846911371538
N	1.36480544924623	-1.78157099459622	3.57937572353366
C	0.89335024384848	-0.55369242263867	1.34349172561463
C	1.99599297030889	-0.38770722177012	0.43273199238257
C	3.26994722291003	-0.94399898522735	0.66985441890791
C	4.27177240406405	-0.80963128673028	-0.27275434950163
C	4.00124603533786	-0.12384049212661	-1.46925883531338
C	2.74371716794172	0.40580240968461	-1.66542897085357
C	1.12226097588165	-1.22637471612640	2.58082662835838
C	-2.08563110105290	0.59091857947522	-0.20949684035891
C	-2.61052806958845	1.76336923979643	-0.76965766130185
C	-3.99051672189510	1.88391660937546	-0.93625350576246
C	-4.84794909000683	0.84964834609562	-0.55148372704366
C	-4.31665615917075	-0.31587508559113	0.01002671347762
C	-2.94012385134687	-0.45022321914242	0.18021677822531
H	3.43384178310612	-1.47817384005900	1.60564594871872
H	5.25934271806105	-1.23453605739038	-0.08805691465457
H	4.76145453132312	0.00859141638670	-2.23788754245406
H	2.46719260904184	0.95804741201776	-2.56173255240586
H	-1.95010938141856	2.58221867131827	-1.04992016334432
H	-4.39638792737377	2.80171285228304	-1.36534043622091
H	-5.92606074279000	0.95076155011084	-0.68956679086634
H	-2.51576578582782	-1.35626013951818	0.61214361329912
H	-4.97653796785675	-1.13197946163225	0.30957134523346

Table B.32. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 BBr_2 hydrazone.

31
Coordinates from ORCA-job BBr2_pNO2

Br	0.74089144965784	2.97263281241779	-0.53169211122051
Br	-0.07296679837307	0.75094959381103	-2.95037651245124
B	0.43090171219741	0.99736620851986	-0.99837737036298
N	-0.63845387164275	0.37938892613098	-0.06460156730257
N	-0.34905165783907	-0.29063283932946	1.03295653951598
N	1.81021714839217	0.27382777011739	-0.72962852675970
N	1.33849522105136	-1.86786657956073	3.54216354364385
C	0.90495512295214	-0.61162726210806	1.31372349739823
C	2.03288420887435	-0.40729551824536	0.43619468925697
C	3.31080296779234	-0.93378348088530	0.70781487177144
C	4.34006810745907	-0.75529526825478	-0.19911527295148
C	4.09082309662181	-0.05445978679324	-1.38988876773404
C	2.82520178117679	0.44373456343933	-1.62002323312929
C	1.11416824198942	-1.30078510319239	2.54643655227529
C	-2.04783943061303	0.53911298372263	-0.26875111814669
C	-2.55838174460070	1.67544684083482	-0.91681856420332
C	-3.93145576819202	1.80970307806190	-1.09671645806363
C	-4.78203682158076	0.80810597451232	-0.62984052866140
C	-4.29170314284987	-0.32581622441447	0.01951069358363
C	-2.92023431601968	-0.45915718075890	0.19772206632473
N	-6.24613595711523	0.95012051413784	-0.82878457355226
H	3.45829803418088	-1.47880689132061	1.64006747925237
H	5.33212546277039	-1.15714121667692	0.01095490397171
H	4.87201072506669	0.11290341158961	-2.13010894189442
H	2.56365839861943	1.00354668691551	-2.51623476422902
H	-1.89025923021476	2.46470254203217	-1.25507436470519
H	-4.35189749158223	2.68683638976374	-1.58724982582606
H	-2.51163371828028	-1.33624325969101	0.69727356243607
H	-4.98646861438919	-1.09032647708492	0.36442420408956
O	-6.64882730833461	1.96533037124100	-1.40733989347691
O	-6.97105497817485	0.04380274806823	-0.40325825184912

Table B.33. PBE/TZVP-optimized Cartesian coordinates of BI_2 hydrazone.

29
Coordinates from ORCA-job BI2

I	0.81296158958630	3.23309102911104	-0.48185895349688
I	-0.23601979535432	0.73037190196394	-3.16325921350230
B	0.38939123414194	1.00661402935720	-0.98317327007516
N	-0.66782564821268	0.42937160130286	-0.02512674243416
N	-0.36887445626818	-0.22432638473700	1.07705921787451
N	1.75978463267054	0.27272715682769	-0.73978039992901
N	1.34689701830281	-1.79099018859664	3.57731222652564
C	0.88799076834632	-0.56404668173492	1.33810874527327
C	1.99439944900315	-0.40116405742593	0.43104692087694
C	3.26639944015732	-0.95533608907570	0.67889419701006
C	4.27977530218103	-0.81141015438653	-0.25017609410108
C	4.02120014682470	-0.11430031424354	-1.44198169933445
C	2.76398630037275	0.41161180045960	-1.64987640891884
C	1.11027568894298	-1.23579760766182	2.57732263219381
C	-2.08760075233957	0.58981361384394	-0.20952057206062
C	-2.61755012478299	1.75384879945814	-0.78119449143878
C	-3.99956897790237	1.87546346843044	-0.93191780464016
C	-4.85400261647359	0.84878536432172	-0.52267825182422
C	-4.31757651672061	-0.31029843969554	0.04753062073827
C	-2.93965869586384	-0.44442947161635	0.20392599141011
H	3.42071925005623	-1.49392715035223	1.61378589513602
H	5.26619263273338	-1.23518974566919	-0.05717461433634
H	4.79004338708879	0.03020019236930	-2.19977505545820
H	2.49985394499880	0.97144909530029	-2.54565932704046
H	-1.96142441308419	2.56986653920139	-1.07972256994029
H	-4.40801785800128	2.78826926519164	-1.36905321430076
H	-5.93356520999314	0.95010532653509	-0.64893039314679
H	-2.51280196250602	-1.34575372652132	0.64291759935551
H	-4.97479175790429	-1.12165017095756	0.36513902458434

Table B.34. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 BI_2 hydrazone.

31
Coordinates from ORCA-job BI2_pNO2

I	0.82447573702815	3.21097439968982	-0.55709139259859
I	-0.17010206577803	0.62827282922447	-3.18879519422178
B	0.42769134846032	0.97618989635624	-1.01111231895251
N	-0.63342097626405	0.40564104379297	-0.04978845235809
N	-0.33822246955602	-0.26090258067393	1.05072378540900
N	1.80625645706035	0.26044586371697	-0.74177128890655
N	1.35083011301463	-1.85875732006029	3.54419757568790
C	0.91152071012159	-0.60610047173859	1.31481103952035
C	2.03148868132023	-0.42382772452918	0.42214306712032
C	3.30409832448235	-0.96570319088079	0.68528090959191
C	4.33133268691700	-0.79689838465026	-0.22606954450495
C	4.08339440087493	-0.08695448015536	-1.41102322495455
C	2.82209188652945	0.42547375768475	-1.63340437847845
C	1.12441736370298	-1.29251033852505	2.54850054442018
C	-2.04660721888316	0.58762351657196	-0.22076483371745
C	-2.55826428099434	1.71843292994850	-0.87643644465063
C	-3.93413345385290	1.87219954391247	-1.01894745769718
C	-4.78610040512722	0.89268280334799	-0.51134338509155
C	-4.29479888050548	-0.23803921694426	0.14280519248183
C	-2.92119965587168	-0.38920866695843	0.28571874791407
N	-6.25352085846631	1.05522276979581	-0.66976257205792
H	3.44971767556611	-1.51331874366997	1.61632899455529
H	5.31969263194365	-1.21107773382437	-0.02278514794952
H	4.86223398077162	0.07736991799045	-2.15443505825785
H	2.56482275498725	0.99276846899042	-2.52644187128795
H	-1.89066212205364	2.49403606848654	-1.24612389344751
H	-4.35395152545500	2.74764489509740	-1.51305648842862
H	-2.51315124159749	-1.26424069822104	0.78902359844691
H	-4.99071060616737	-0.98629385100537	0.51949057724988
O	-6.65686756706051	2.06729296628376	-1.25296112588381
O	-6.97932508514739	0.16758307494632	-0.20817605695218

Table B.35. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 BCl_2 hydrazone.

31
Coordinates from ORCA-job pNO2BCl2

Cl	-0.99037691352202	1.31867105461115	-1.37573379523742
Cl	-1.71149831795479	-1.52303605301806	-2.36759156434753
B	-1.11260107370936	-0.49920264575948	-0.93820536883705
N	-2.01539427092610	-0.69633240204598	0.31510736037127
N	-1.55026305655057	-0.77428444281847	1.54420535897288
N	0.35577394054151	-0.96700600036888	-0.55847299389609
N	0.53894017468657	-0.93242860827095	4.23717797274469
C	-0.24638911638848	-0.86179791898038	1.76536549639348
C	0.76283748574269	-1.02522603912305	0.74597604498865
C	2.10967917458640	-1.31038467522748	1.04678157429374
C	3.01571842704641	-1.52628729071844	0.02415885572497
C	2.57555854275164	-1.46704874713098	-1.30826026452057
C	1.24864257881398	-1.18560148228234	-1.56067449137952
C	0.15612874048285	-0.89694852005858	3.13469214014514
C	-3.44431698952324	-0.72366313282391	0.24560076751313
C	-4.12675384097521	-0.06020443162761	-0.78761333041790
C	-5.51675905184816	-0.09722291364748	-0.83302961192322
C	-6.21402686504178	-0.79500852406914	0.15273428502247
C	-5.55225614997074	-1.45349779219534	1.19024619691224
C	-4.16388714423106	-1.41964612496203	1.23266271989850
N	-7.69621330172646	-0.83834103672656	0.09789450658440
H	2.40824406461894	-1.35974788201978	2.09399898847090
H	4.05974011173761	-1.74354252748150	0.25425193156775
H	3.25504051223111	-1.63037874959899	-2.14332671324169
H	0.83889954095504	-1.11580609162771	-2.56669950348552
H	-3.57794273312224	0.50439212605353	-1.53842041528894
H	-3.62314221865905	-1.92972591421681	2.02835714186094
H	-6.06851741901307	0.41513620659643	-1.62013903470887
H	-6.13044833930481	-1.99170713429678	1.94043329785923
O	-8.28443604792094	-1.43809504857298	1.00480337809556
O	-8.25045845080667	-0.27278210259140	-0.85116224513563

Table B.36. PBE/TZVP-optimized Cartesian coordinates of $AlBr_2$ hydrazone.

29
Coordinates from ORCA-job AlBr2

Br	0.55487032151866	3.42384182975010	-1.13158014881917
Br	0.03686073666205	0.48100042617593	-3.60360787867467
Al	0.32537797059603	1.15316538837794	-1.42122037673918
N	-0.82468775728265	0.29157017484388	-0.14202086445990
N	-0.41839908921496	-0.53020349171457	0.79751367208014
N	1.98229086522176	0.33503340619633	-0.79695283019974
N	1.17453714155409	-2.52900944419704	2.98279798796683
C	0.84734276987798	-0.88993731210744	0.98919470679536
C	2.05522371752139	-0.52304796553058	0.26856642459985
C	3.31298884581296	-1.04582628063567	0.65134682290182
C	4.45573862487249	-0.69451964744682	-0.04369532483008
C	4.35950167032872	0.18482003093979	-1.13489040331526
C	3.11332248443487	0.67120252556022	-1.47367715177985
C	1.00919159783673	-1.79292108090986	2.09078577393696
C	-2.23753230899743	0.51337617561435	-0.15249759104446
C	-2.76211519525009	1.42301097415859	-1.08349756545259
C	-4.13400614478200	1.67113064996981	-1.12816316655314
C	-4.99702633551360	1.01549826067022	-0.24787740467036
C	-4.47289237740643	0.10574733850444	0.67837193791895
C	-3.10477268750499	-0.14957522356133	0.73226045941280
H	3.35851901454398	-1.72575657184051	1.50187516878661
H	5.42283544901276	-1.10019532042109	0.25671585608413
H	5.23403998623156	0.48637555582736	-1.70926805780841
H	2.97426474558862	1.35701897044733	-2.31166679559278
H	-2.10845304795171	1.95698041863586	-1.77911180376669
H	-4.52532238283038	2.38365496851758	-1.85633312930114
H	-6.07026427725777	1.21096813208888	-0.28232319584144
H	-2.69223221245178	-0.85586832623151	1.45143522515755
H	-5.13861012517084	-0.41456555768221	1.36970364920784

Table B.37. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 $AlBr_2$ hydrazone

31
Coordinates from ORCA-job pNO2_AlBr2

Br	0.53106410148590	3.41262753621780	-1.13794846185317
Br	0.05124119037311	0.45826337258200	-3.60750428812188
Al	0.33530597400309	1.14543681939963	-1.43500737186471
N	-0.81781968806537	0.27223445745972	-0.16368339130959
N	-0.41002624907250	-0.56504680194940	0.77018333361351
N	1.99799970888696	0.34198516819724	-0.79572197506734
N	1.17903619818290	-2.59369228037210	2.92592709408822
C	0.85177896564646	-0.92140290999900	0.95925027371068
C	2.06778764244204	-0.53394715434122	0.25326526118020
C	3.32263975537083	-1.05388132641739	0.64068714090286
C	4.47128374201782	-0.67965520755647	-0.03574368832709
C	4.37950074235158	0.21928291329232	-1.10873403074824
C	3.13178170530340	0.70206879471241	-1.45276002583299
C	1.01371480289539	-1.84363465445376	2.04647632940101
C	-2.22451443764879	0.48914286646222	-0.16183902907981
C	-2.76543541991696	1.38226543686493	-1.10440591804806
C	-4.13316927786327	1.62814012435059	-1.14111368994696
C	-4.96216409483300	0.97439641425125	-0.22994112314900
C	-4.44645886090166	0.08238699121138	0.71400921809367
C	-3.07989109695306	-0.16095574647298	0.74865491510006
N	-6.42018855111537	1.23386911895112	-0.26351160472301
H	3.36613475593623	-1.74941965543541	1.47865269919601
H	5.43816240953703	-1.08358109483006	0.26743745075358
H	5.25769686173575	0.53963363050732	-1.66711008500164
H	2.99643253297506	1.40355696597032	-2.27834178767823
H	-2.12486779190588	1.90266171816370	-1.82085682167897
H	-4.56563460631931	2.31853856209863	-1.86445766709975
H	-2.65718388545072	-0.85169346509688	1.47608554336824
H	-5.12489432627581	-0.41033823611230	1.40959803582779
O	-7.12547714503565	0.63855142530791	0.55965883113059
O	-6.84112007978621	2.02804459103646	-1.11244233783596

Table B.38. PBE/TZVP-optimized Cartesian coordinates of AlI_2 hydrazone.

29
Coordinates from ORCA-job AlI2

I	0.61079626024205	3.66717266204609	-1.06406808854160
I	0.00877825927124	0.43527557023127	-3.82376982077915
Al	0.33066401803864	1.16628766681022	-1.41525144889933
N	-0.82434963683547	0.29956617709533	-0.13802557678595
N	-0.41682583878364	-0.51750702585414	0.80530599489694
N	1.98788830376871	0.33355418752999	-0.79142695530549
N	1.16838349454236	-2.50319984458454	3.00669218727662
C	0.84732164990336	-0.87758703900905	1.00125769677115
C	2.05644555933475	-0.52022806658589	0.27771821146262
C	3.31074302414692	-1.05037262303873	0.66137501530648
C	4.45533561866035	-0.71114021337268	-0.03673970495937
C	4.36354564169511	0.16337265714786	-1.13185685178279
C	3.12053118327014	0.65769248021918	-1.47094106005458
C	1.00608194723952	-1.77323048431397	2.10943613930549
C	-2.23909502695179	0.51184664860301	-0.15047865021260
C	-2.76924689137534	1.43374923165927	-1.06573304540262
C	-4.14330776846367	1.67026903584042	-1.11083245657132
C	-5.00337297818567	0.99037724084923	-0.24644220266762
C	-4.47395138242561	0.06840357261312	0.66464407459238
C	-3.10376968224221	-0.17540439535668	0.71837876256382
H	3.35272473028996	-1.72658147731064	1.51502765213132
H	5.41983746935974	-1.12250015997273	0.26423773554767
H	5.23928899021537	0.45594598192506	-1.70906035574108
H	2.98627621231641	1.34104663737312	-2.31183509992729
H	-2.11841411799760	1.98939155793228	-1.74652675648943
H	-4.53832779332045	2.39316773575748	-1.82664833122324
H	-6.07830601670873	1.17631296902500	-0.28146943349256
H	-2.68795615214963	-0.89165063939399	1.42561501469500
H	-5.13712707985481	-0.47106104186488	1.34360135028655

Table B.39. PBE/TZVP-optimized Cartesian coordinates of para-NO_2 AlI_2 hydrazone.

31
Coordinates from ORCA-job pNO2_AlI2

I	0.56959082967763	3.63309941561351	-1.10513157624536
I	0.02859459861660	0.35088897960014	-3.82966893284448
Al	0.34586269378840	1.13029799077344	-1.44023555899469
N	-0.80251422535541	0.26592908918340	-0.15125358820362
N	-0.38462806858022	-0.54346748424870	0.80235765094092
N	2.01397957110716	0.32472785013070	-0.79924585591907
N	1.21770629508996	-2.50146702894176	3.01153204218556
C	0.87857792644848	-0.88968443438652	0.99684038908490
C	2.08981029022191	-0.52294232179275	0.27208365016624
C	3.34564746916299	-1.03759954065556	0.66325838946435
C	4.49030555550554	-0.68711978702299	-0.03235101208273
C	4.39244193861565	0.18204815458728	-1.12892469258315
C	3.14390472682495	0.66092736905419	-1.47523479857858
C	1.04750258603854	-1.77850398933441	2.11057733430415
C	-2.21182135999540	0.46973872242396	-0.14576698909322
C	-2.76898298391633	1.33822548584475	-1.10161093379305
C	-4.13952109329381	1.56895820436736	-1.13401056230089
C	-4.95550770732803	0.92545184774196	-0.20427326870406
C	-4.42383214925736	0.05831357371540	0.75403705861688
C	-3.05466098268524	-0.17031569028079	0.78358791341400
N	-6.41630871327987	1.16862237324152	-0.23282293184232
H	3.39270801591791	-1.71025885355716	1.51946481945354
H	5.45809298136593	-1.08668884196293	0.27369642425429
H	5.26681379580416	0.48328248539646	-1.70377631248268
H	3.00467873589368	1.33984476483606	-2.31897178568701
H	-2.13912718142479	1.85049822013721	-1.83296843822603
H	-4.58378607149053	2.23993877224442	-1.86843048543380
H	-2.62020464472616	-0.84174544249321	1.52199443896107
H	-5.09192327018319	-0.42673413402385	1.46483127870100
O	-6.85233642141790	1.93741258813109	-1.09735278572184
O	-7.10888094614527	0.58667870667779	0.61051720218968

CHAPTER C. SUPPORTING INFORMATION FOR CHAPTER 4

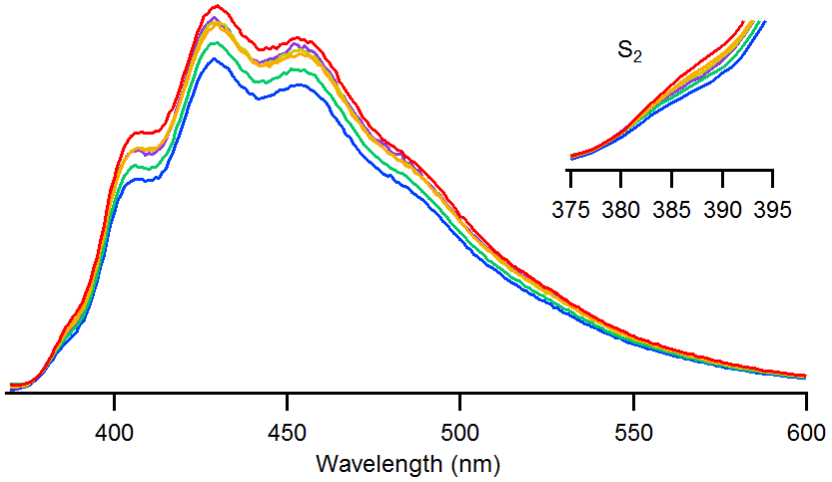


Figure C.1. Emission spectrum of DPH in EtOH, EG and Gly. Sample was excited at 350 nm. Traces are the average of 5 scans. An integration time of 0.1 s, excitation slit widths of 0.5 nm, emission slit widths of 5 nm, and a step size of 1 nm were used. Purple trace is 0 % Gly, blue trace is 10 % Gly, green trace is 20 % Gly, yellow trace is 30 % Gly, orange trace is 40 % Gly, and the red trace is 50 % Gly. Spectra were collected by Adam Petrucci.

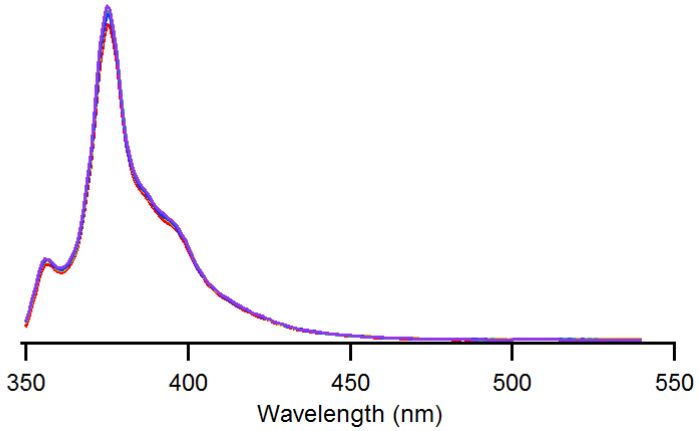


Figure C.2. Emission spectrum of Az in EtOH, EG and Gly. Sample was excited at 326 nm. Traces are the average of 5 scans. An integration time of 0.1 s, slit widths of 2 nm, and a step size of 1 nm were used. A is 0% Gly and F is 50% Gly. Spectra were collected by Adam Petrucci.

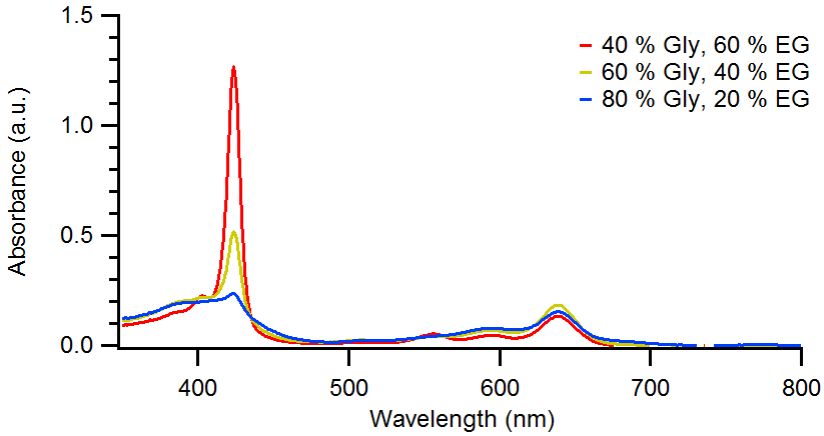


Figure C.3. Observed π-stacking of Zn TPP in viscous solutions after 24 hr incubation period. The intensity of the Soret (420 nm) decreases upon π-stacking.

Table C.1. Solutions to probe viscosity sensitivity of potential SOKR fluorophores with a viscosity range of 20-200 cP.

PS-EG (%)	**EG** (%)	**Gly** (%)
10	90	0
10	80	10
10	70	20
10	60	30
10	50	40
10	40	50

CHAPTER D. APPENDIX FOR CHAPTER 5

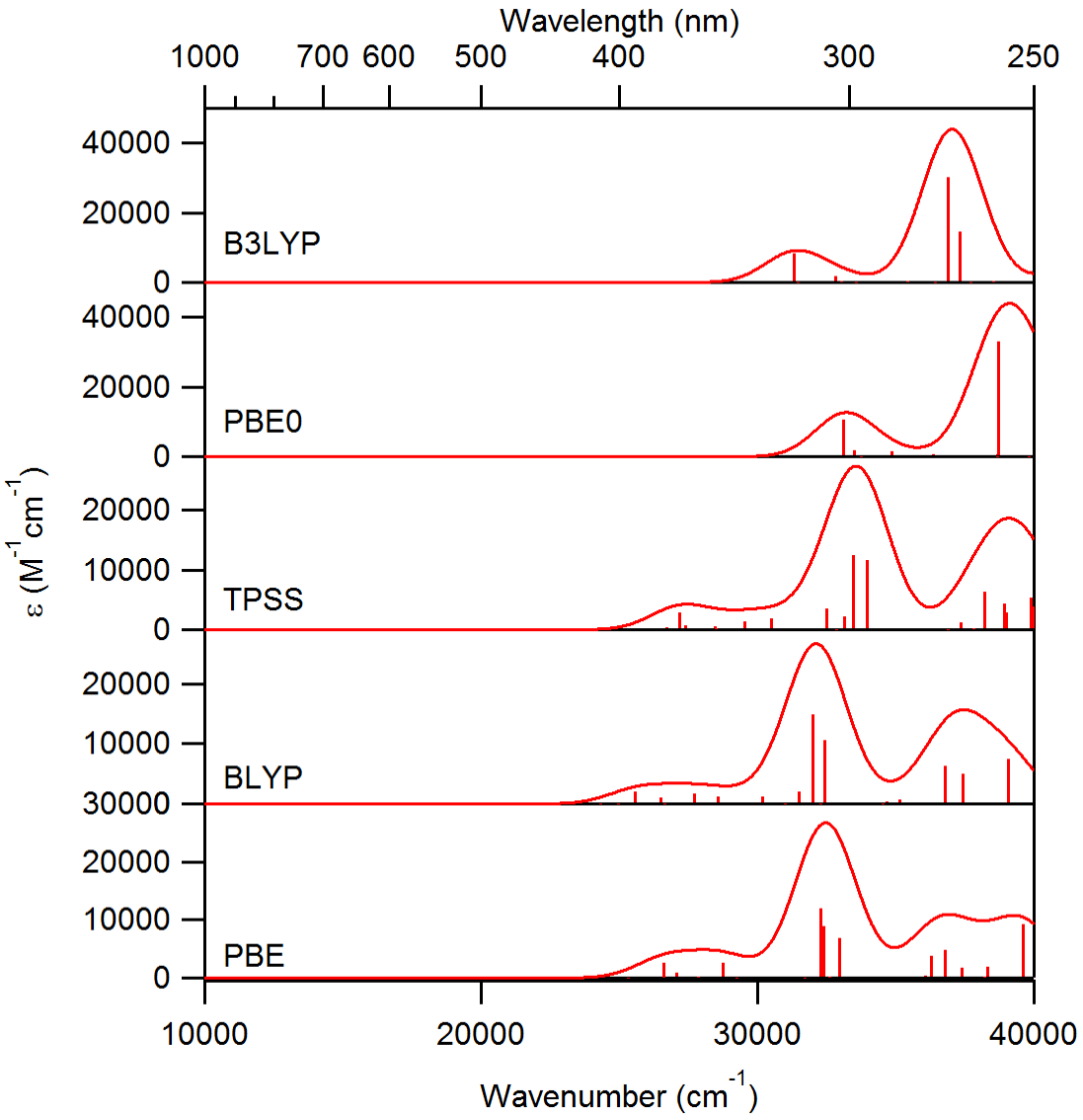


Figure D.1. TDDFT-predicted Abs spectra of TOQ monomer with series of functionals. Spectra were plotted in the orca_mapspc utility program with a full width at half maxima of 2500 cm^{-1}.

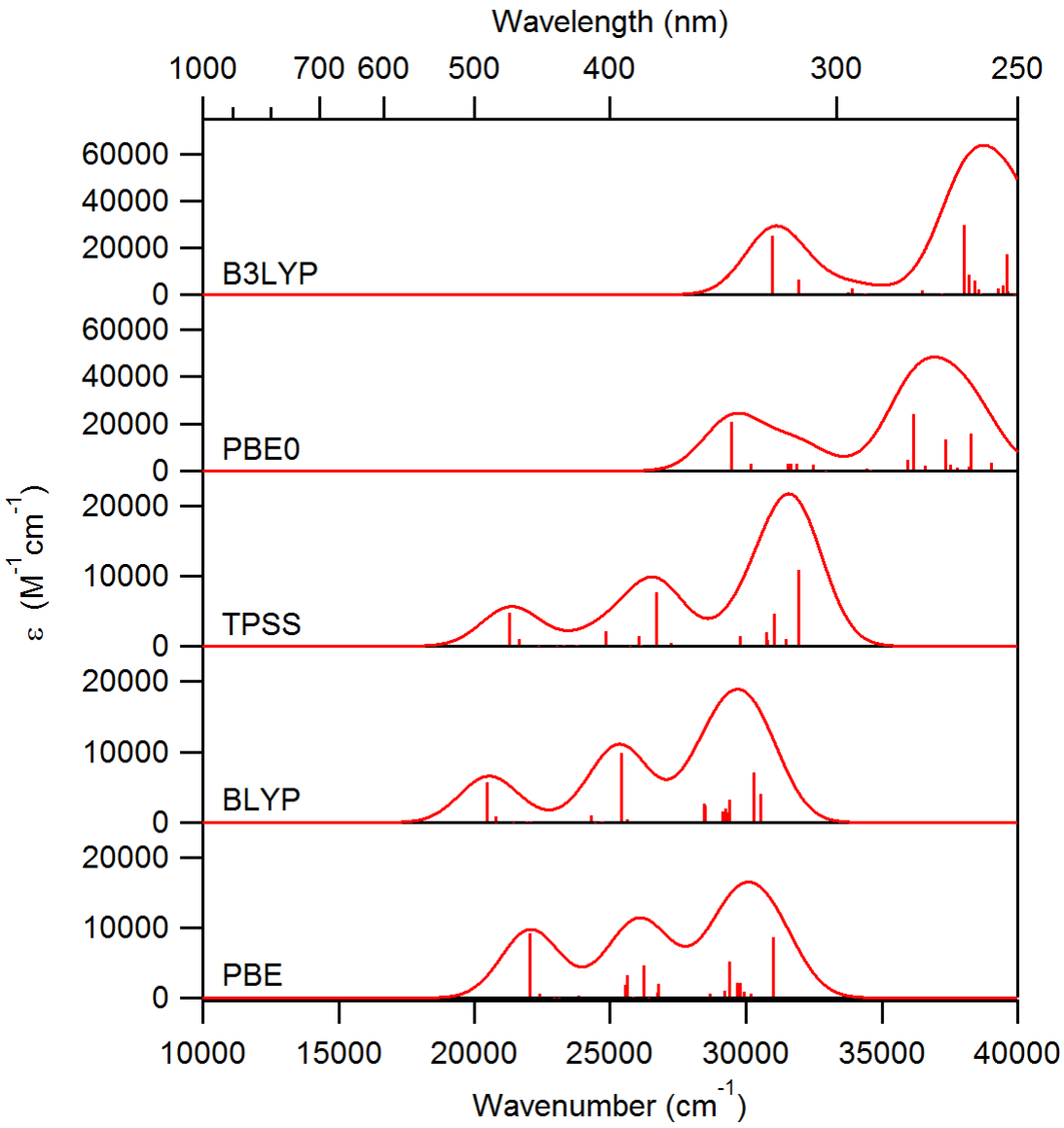


Figure D.2. TDDFT-predicted Abs spectra of TOQ dimer with series of functionals. Spectra were plotted in the orca_mapspc utility program with a full width at half maxima of 2500 cm^{-1}.

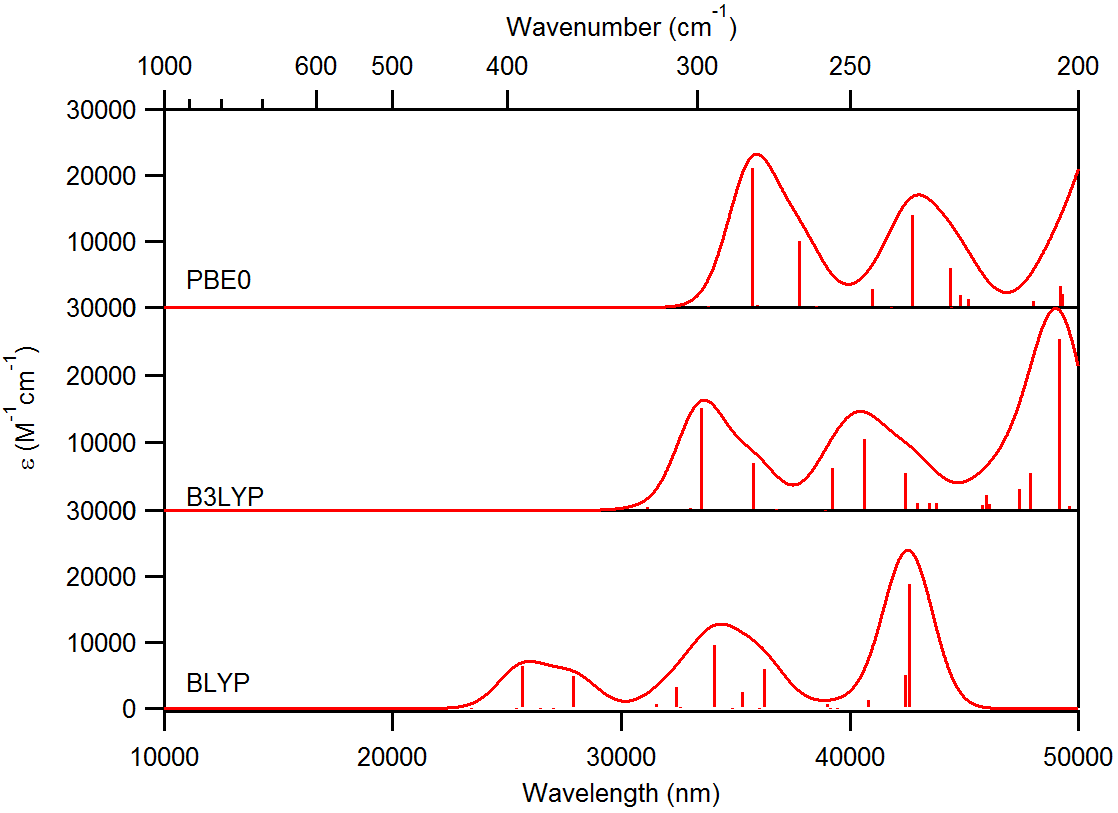


Figure D.3. TDDFT-predicted Abs spectra of TOP monomer with series of functionals. Spectra were plotted using the orca_mapspc utility program with a full width at half maxima of 2500 cm^{-1}.

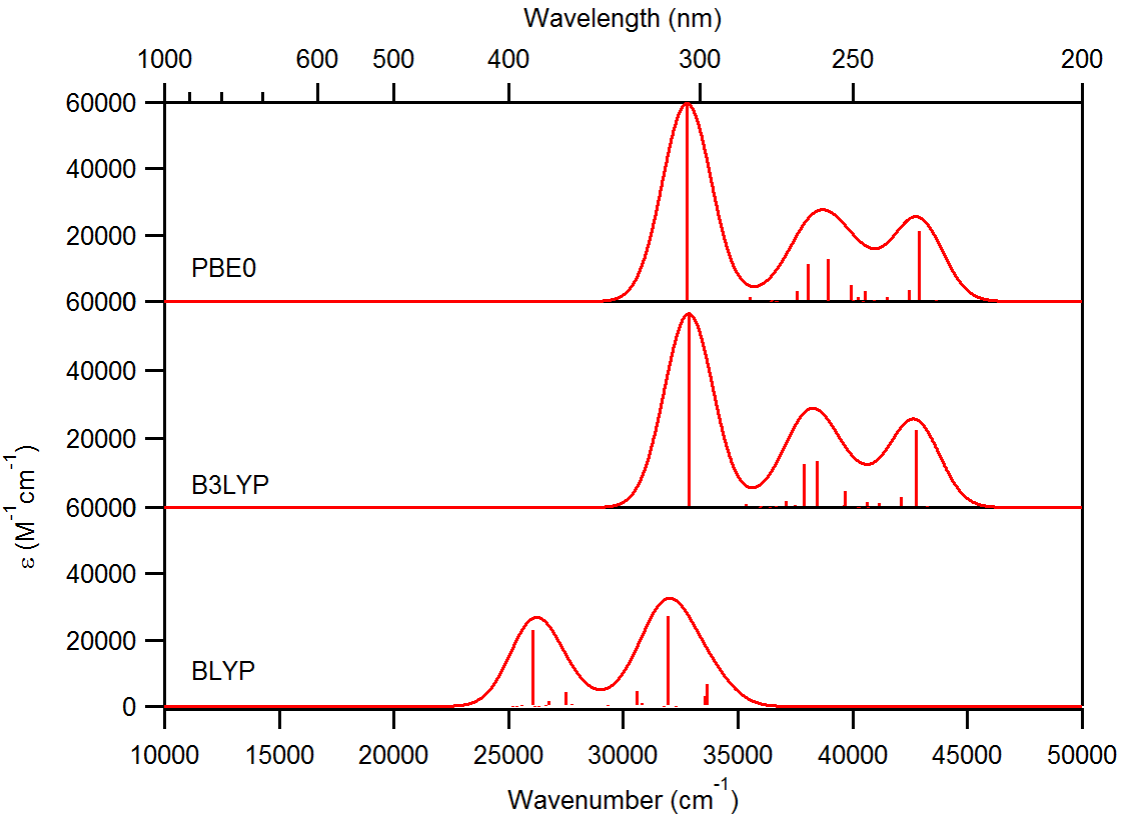


Figure D.4. TDDFT-predicted Abs spectra of TOP dimer with series of functionals. Spectra were plotted in the orca_mapspc utility program with a full width at half maxima of 2500 cm^{-1}.

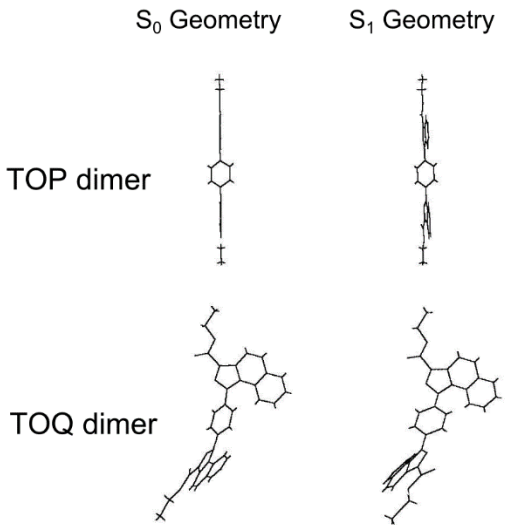


Figure D.5. Different viewing angles to highly the minute differences in the PBE0-predicted ground state and excited state geometries of the TOP and TOQ dimer.

Table D.1. PBE0/TZVP-optimized Cartesian coordinates of TOP dimer in water.

56
Coordinates from ORCA-job GeoOpt2_PBE0

O	13.55181667680587	6.39042252814182	4.40636973089533
O	11.54402043751957	7.30920452077161	4.84370059189358
N	10.37942423517010	5.04692964027112	3.66095472351201
N	10.16284348229461	3.89870232024539	3.07938733384632
N	11.33105439999813	3.28294088794191	2.76758440099413
C	11.68031851631323	5.20935277541991	3.76002956845143
C	12.34573827863733	4.08908337254129	3.19931435412774
C	13.66249438605798	3.67696974722365	2.99555168196478
C	13.87673201670564	2.48246319445025	2.35562234290611
C	12.79523700676744	1.70180859819825	1.91021470713765
C	11.51000394442657	2.10527677670207	2.11755873020693
C	12.23250829661344	6.42582278196019	4.39914837770270
C	14.23323620982552	7.51899089621730	5.01321711699206
C	15.71151433881782	7.28600140121469	4.87778685074230
C	8.86836759125778	3.38745567728826	2.78796514003858
C	8.00222960989696	4.18714086113498	2.05760959046675
C	6.72476277578583	3.72345977466272	1.80119137226448
C	6.35752667110161	2.47622398484133	2.28410337299688
C	7.22045008768309	1.68018279899747	3.02387863864439
C	8.49843308701541	2.14370006092680	3.27987838887352
H	14.47080932356704	4.30576995962308	3.33907762250553
H	14.88807808303548	2.13738389016112	2.18208141671052
H	12.96315175585576	0.76656720600514	1.39353843808498
H	10.63117082228960	1.56571001673166	1.79700300507172
H	13.92049460685076	7.57898873087532	6.05709743584611
H	13.91082259299530	8.42549533685195	4.49840099912109
H	16.01570759269542	6.37025208660502	5.38835620359413
H	16.24446649582536	8.12312826529381	5.33338724531553
H	16.00640081991497	7.22428665512941	3.82857568571704
H	8.32250327851451	5.15309626103601	1.68919839796612
H	6.02584331375371	4.32019926900472	1.22876939697967
H	6.89748460085031	0.73162944098733	3.43480208625950
H	9.17434181142545	1.55711753885981	3.88997776759057
O	0.63293702921292	1.67982507947924	1.89336235346364
O	1.33721201955278	3.63557592967540	2.75719299665108
N	3.96449455349219	2.72598094173791	2.37056170785348
N	5.02574317893819	2.03631639502523	2.05131079074888
N	4.68431597783153	0.86164336041936	1.46420610965314
C	2.90838601621437	2.02563091477684	2.02043917127403
C	3.31998475201899	0.80256605367230	1.43171438643187

C	2.71000157488525	-0.31837060361105	0.86872324193700
C	3.51312648667390	-1.30718968528715	0.35955524260705
C	4.91351916357333	-1.18460532777215	0.39603860638421
C	5.50530876685549	-0.08995203192185	0.95225439444019
C	1.54734759760909	2.55327961493081	2.27096826075415
C	-0.75463116785404	2.05947419496413	2.08397507859429
C	-1.60823220228982	0.91537276652344	1.61379862334382
H	1.63119077990705	-0.37499671292272	0.84354603269756
H	3.06803256920821	-2.18874440214616	-0.08426537647445
H	5.54779938702083	-1.95623852700770	-0.01933091907300
H	6.57084374864460	0.07787897663346	1.00596122694522
H	-0.93650248316156	2.97078902989633	1.51196312150157
H	-0.90324770320464	2.27522874642674	3.14321255885154
H	-1.44170489121934	0.70767073087785	0.55503359719723
H	-2.65940386729187	1.18024741488895	1.74846218922392
H	-1.40788559288520	0.01062009042533	2.19084169257335

Table D.2. BLYP/TZVP-optimized Cartesian coordinates of TOP dimer in water.

56
Coordinates from ORCA-job DimerWaterOpt

O	13.62970681791891	6.43984191307163	4.43721822690484
O	11.57315297452956	7.34619821666788	4.88262001948770
N	10.42091791306731	5.05780621673040	3.68369142236986
N	10.20107937208661	3.88851802746669	3.08732316814096
N	11.40633054383108	3.27051622826152	2.76035563184471
C	11.74176669610697	5.22849594419285	3.77645604953479
C	12.42331894709130	4.10689775287520	3.20310556719242
C	13.75425855183698	3.70644717542254	2.99146584869444
C	13.99077541270091	2.50514469386966	2.33986125992554
C	12.91459230641259	1.70077733351067	1.89033484112303
C	11.60801147632303	2.08426292823570	2.10074945395869
C	12.28193646028870	6.45989113882452	4.42774041206263
C	14.31453837311269	7.60359537050430	5.06874457729066
C	15.81009045777388	7.38666664315957	4.92487725646506
C	8.89226620852837	3.36805506508651	2.78773404138670
C	8.02082377749596	4.16525498240151	2.03404374488000
C	6.73185432358039	3.69659146440379	1.77616108904139
C	6.35278943094389	2.44508687702457	2.28015036929586
C	7.22185760374412	1.65071042429014	3.04056120835312
C	8.51163958983769	2.11928764642898	3.29790596167241
H	14.55854018384074	4.34678046485061	3.33685011168310

H	15.01226889173442	2.17695107031732	2.16394134908043
H	13.09936119346896	0.76542674981464	1.36953066649832
H	10.73578514532237	1.52672548864453	1.77932313605496
H	13.99643837473600	7.64234380436150	6.11657823181889
H	13.97110079125207	8.50801412335078	4.55465722782563
H	16.13129120582062	6.46712029721462	5.42995923540963
H	16.33374374577678	8.23222529330142	5.39059278788577
H	16.10666375611526	7.34050881229771	3.86971269566818
H	8.34372279421601	5.12909671792270	1.65009001361885
H	6.03645323900471	4.29113774521724	1.18906691589926
H	6.89512660437219	0.70497403331145	3.46549783641907
H	9.18513991524549	1.53633233133992	3.92142426545934
O	0.53670623494305	1.66982649457207	1.88420750228682
O	1.28172111733210	3.65423090611815	2.75442198415480
N	3.92113819211278	2.70777187196528	2.37230436110027
N	5.00253970878111	2.00203130697575	2.04978077247671
N	4.64108083874827	0.80118955863400	1.44241913338076
C	2.84768616000834	2.00142764126379	2.01026932638793
C	3.25256150241876	0.76463820564729	1.41254467494165
C	2.62787606668829	-0.35798761693709	0.84140264809678
C	3.42304784775724	-1.36844673027534	0.32184895593472
C	4.83581803661224	-1.26827060711426	0.35736750304151
C	5.45426622232597	-0.17369874778940	0.92145385175885
C	1.48201865653519	2.55203015358574	2.26483907065144
C	-0.88339482394589	2.07683170497621	2.08178475185592
C	-1.75598406583471	0.92517345168744	1.61499108646142
H	1.54413740839192	-0.40234873084045	0.81515525561945
H	2.96074259871868	-2.24510425766621	-0.12550700235132
H	5.45940954729934	-2.05282403080155	-0.06260264122258
H	6.52663933832737	-0.02453281625331	0.97611457390172
H	-1.04828173856729	2.99003750690632	1.49950941674305
H	-1.01788768738697	2.29899569022869	3.14604999999717
H	-1.59707309003086	0.70936961543773	0.55103864941082
H	-2.80938206589595	1.20490418128006	1.75359893037893
H	-1.56410624345564	0.01683248002601	2.19954037504691

Table D.3. B3LYP/TZVP-optimized Cartesian coordinates of TOP dimer in water.

56
Coordinates from ORCA-job GeoOpt2_B3LYP

O	13.57856442750396	6.41222879679023	4.42445451356609
O	11.55463503008940	7.31567653763235	4.87128558672060
N	10.39843303817521	5.04816356914787	3.67412520558462
N	10.17931416075080	3.89555308962350	3.07993017136680
N	11.36286728999093	3.28477014370788	2.75745484719163
C	11.70323102223509	5.21722161065050	3.76848019651644
C	12.37626962282127	4.10159176988475	3.19459379484798
C	13.69816448163200	3.69903237098940	2.98343759826334
C	13.92291690012981	2.50685562262878	2.33436463677648
C	12.84552138688352	1.71553477315693	1.88748268709873
C	11.55339230927267	2.10651368927808	2.09985506250146
C	12.25086480139329	6.43672262706810	4.41779548842990
C	14.26822433655673	7.54943454548055	5.04469959205389
C	15.75264810737047	7.31627195609854	4.90272658343984
C	8.87729783865688	3.37903654193919	2.78708359126656
C	8.01886547055901	4.16020491382974	2.02135338940881
C	6.73803068361047	3.69352471057392	1.76668689990157
C	6.35690989978411	2.46115828812978	2.28644209622891
C	7.21325963835941	1.68249339733195	3.05831284703755
C	8.49467467995885	2.14894734088720	3.31241156572472
H	14.50190565943134	4.33069890974265	3.32777859353050
H	14.93602962454767	2.17285988873546	2.15676581373887
H	13.02200042549037	0.78488575843788	1.36768008821092
H	10.68114163151672	1.55868004733240	1.78086982704279
H	13.95609924562441	7.59736157081012	6.08781537317156
H	13.94200677730729	8.45558998543224	4.53487880865060
H	16.05866290582682	6.39744957175698	5.40573948970937
H	16.28557653164441	8.15047051276928	5.36343734728397
H	16.04510288658297	7.26279519705252	3.85285796848619
H	8.34522239874259	5.11193690994028	1.62609384980493
H	6.04964457827816	4.27625574841396	1.17006068447150
H	6.88417842942820	0.74741769770795	3.49086147717966
H	9.16257760289225	1.57632376630108	3.94179044690840
O	0.59439339533691	1.68121629963537	1.89304506830978
O	1.31778697472319	3.63752185262754	2.76622599949360
N	3.94638146911230	2.71218893778860	2.37789800844828
N	5.01617610106678	2.01815107439612	2.05598317481482
N	4.66412233747210	0.83552759909086	1.45969117033943
C	2.88481047471431	2.01438823593379	2.02295767107391
C	3.29218307887430	0.78689747895290	1.42803671255077

C	2.67487822891046	-0.33163868621533	0.86051454436278
C	3.47193674257802	-1.32852007855170	0.34698771950479
C	4.87660947586121	-1.21861582397648	0.38509619959666
C	5.47960593995938	-0.12740644472624	0.94498149708238
C	1.52192492397721	2.55110793479440	2.27546795271007
C	-0.80878414195059	2.06620255243230	2.08276562516206
C	-1.66478134047606	0.91802666505624	1.60546267738846
H	1.59730651593290	-0.38182734505800	0.83452616488673
H	3.01978794349641	-2.20421434205843	-0.09847838274725
H	5.50284219058734	-1.99529021273448	-0.03020605456030
H	6.54516440837708	0.02827033885722	1.00173208293651
H	-0.98366740187395	2.97718578667130	1.51088685975750
H	-0.95559502525997	2.27811259482440	3.14148638576488
H	-1.49630522778166	0.71290367537384	0.54694102830779
H	-2.71582059901033	1.18378420494232	1.73861920128553
H	-1.46656744767450	0.01212604868142	2.18063037441559

Table D.4. PBE0/TZVP-optimized Cartesian coordinates of TOP monomer in water.

34
Coordinates from ORCA-job MonoWaterOpt

O	13.50568747574460	6.35143498227784	4.63448798328921
O	11.46997639014030	7.10887273569450	5.22447298866641
N	10.37680975920542	5.01966423799998	3.69516815064052
N	10.19850761262492	3.97958273635925	2.93128852440043
N	11.38743244880734	3.47346420487394	2.51978942064561
C	11.67385045340671	5.21689398561212	3.81616061896505
C	12.37602464056756	4.23422505081721	3.07497719009061
C	13.70522878374471	3.90783598145304	2.79984869842728
C	13.95329371379614	2.84220093136193	1.97290155785748
C	12.89466080717887	2.10378182850494	1.41211617639143
C	11.59893234812544	2.42169503913817	1.68833052665418
C	12.18491088834192	6.33417633780802	4.63970937049628
C	14.15031833756754	7.39955724865737	5.40231929973011
C	15.63556673300194	7.22232529065057	5.25768179418352
C	8.92840600292139	3.46642798282323	2.53231611987759
C	8.09957235937675	4.27952926069175	1.77373503030701
C	6.85419889016070	3.79172093158662	1.40776927240195
C	6.46647084907820	2.51199735248393	1.78698662659844
C	7.31483816172369	1.71285587381516	2.54438497747795
C	8.55772770794334	2.18994240391042	2.93249130351528
H	14.49508569983719	4.50010405121093	3.23790561156832

H	14.97398582059560	2.56607704145458	1.74079897030549
H	13.09098363191224	1.27035761199167	0.75094208392881
H	10.73139652064979	1.91082991527407	1.29698185600781
H	13.82491465461175	7.30851713177898	6.44019435922371
H	13.81139364077028	8.36097585908731	5.01248436626015
H	15.95646736802911	6.25230921990456	5.64212407067784
H	16.14255149497797	8.00069258508387	5.83135468862441
H	15.94408958804573	7.31194595021170	4.21439803553874
H	8.42508983760125	5.27007749371472	1.48071522684771
H	6.18835566831957	4.41230766518291	0.81974619156459
H	5.49371879004264	2.13562427655382	1.49274934818236
H	7.00606891661157	0.71961962083444	2.84669261651638
H	9.21913005253780	1.58756119719640	3.54469895413734

Table D.5. BLYP/TZVP-optimized Cartesian coordinates of TOP monomer in water.

34
Coordinates from ORCA-job MonoWaterOpt

O	13.53985836163884	6.36622345052083	4.65488688987204
O	11.46197681118780	7.12526145323481	5.25087660982400
N	10.36501522735687	5.01544391019954	3.70994320106847
N	10.17483392074034	3.95882459375039	2.92758233583554
N	11.39715371252678	3.45025742928675	2.49313312505001
C	11.68296380276544	5.21628963851909	3.82066117725562
C	12.39333144579886	4.23266814445631	3.06177363930201
C	13.73335215994769	3.91290964288120	2.77534792289820
C	13.99517542202071	2.84289584919711	1.93324433768901
C	12.93675593285755	2.08916454707728	1.36633959019359
C	11.62308186774137	2.39302609936523	1.64802692365123
C	12.19083602791393	6.34324318530803	4.65631482937059
C	14.19594789171598	7.44276365934003	5.44761540022272
C	15.69688007625024	7.26141973056762	5.30710640976373
C	8.88540361406032	3.44495984304403	2.52437011687679
C	8.05865652815977	4.25967949554410	1.73937360420443
C	6.79740632425895	3.77386070861745	1.37574024027785
C	6.38815615440336	2.49560569048861	1.78414219471716
C	7.23294104665291	1.69525804328869	2.56756652478457
C	8.49291997753854	2.16863331851642	2.95213919163297
H	14.52427070501227	4.51078310562794	3.21451828626121
H	15.02339065948050	2.57905587377511	1.69788105213664
H	13.14302758586808	1.25707536783393	0.69840961065716
H	10.75937754899195	1.87010297776356	1.25307638405382

H	13.85828536247642	7.34296121634303	6.48546883821257
H	13.85173392672183	8.40344294770287	5.04820565699669
H	16.01985445104685	6.28894789238124	5.69937230047618
H	16.20181367149978	8.04724254012555	5.88430102223338
H	16.01287383217746	7.34962687200300	4.26010799603423
H	8.39556022851947	5.24478439337050	1.42508770378757
H	6.13811963847301	4.39351159443452	0.77023266376646
H	5.40698123994264	2.12387932594283	1.49382553207715
H	6.91015485575993	0.70794101682367	2.89164802895791
H	9.14755603649357	1.56744045866872	3.58040266985850

Table D6. B3LYP/TZVP-optimized Cartesian coordinates of TOP monomer in water.

34
Coordinates from ORCA-job MonoWaterOpt

O	13.51489067318748	6.35518815943647	4.65108537794268
O	11.46651392010847	7.10097958553028	5.25063744869976
N	10.37399651246622	5.01463869242577	3.70432859335342
N	10.18901743225592	3.97506212374533	2.92452800998126
N	11.39111360171438	3.47443472650347	2.49962515128673
C	11.67528488852942	5.21553458517742	3.82041006251340
C	12.38098218233886	4.24044023973585	3.06261996846501
C	13.71360854661168	3.92261077728561	2.77827908421896
C	13.96798005484411	2.86279965340092	1.93933822171601
C	12.91108343130888	2.11876586940614	1.37464240992928
C	11.61017471643828	2.42590800147649	1.65686076141966
C	12.18603907492074	6.33100321648300	4.65595134948499
C	14.17154497176171	7.40785165209113	5.43285229993122
C	15.66230838700482	7.22449826974168	5.28255043065815
C	8.90906592043181	3.45770398118184	2.52878375435759
C	8.09609111505893	4.24686070275060	1.72227076286814
C	6.84575493897335	3.75671291755860	1.36221095499423
C	6.43562544963270	2.49775871390475	1.79631003833868
C	7.26750483694993	1.72177227694553	2.60103810396069
C	8.51562327501402	2.20186419052167	2.98166001164691
H	14.50128371858219	4.51407045764659	3.21741668018271
H	14.98894533556822	2.59646546222197	1.70262519803582
H	13.11261961773302	1.29223189052879	0.70823445568816
H	10.74807228418666	1.91039850071406	1.26438143953313
H	13.84627771026600	7.30744655600116	6.46830603614597
H	13.83336740613639	8.36937400413386	5.04683190534130
H	15.98077019972829	6.25197031477897	5.66138907352860

H	16.17263618916500	7.99840983149483	5.85895186799031
H	15.96855074619637	7.31827771406998	4.23942866459226
H	8.43486451432765	5.21863726491477	1.38916497860298
H	6.19534794200723	4.35753941121667	0.73942277323612
H	5.46229572133285	2.12105474437288	1.50859107279522
H	6.94330329515220	0.74792342625281	2.94356150632425
H	9.16310743806621	1.61899610235009	3.62443356223640

Table D.7. PBE/TZVP-optimized Cartesian coordinates of TOQ dimer in water.

68
Coordinates from ORCA-job GeoOpt_PBE

O	13.55000236911159	5.75487394021411	5.14307007837661
O	11.50744001290864	6.34875556355827	5.96938776209233
N	10.34646268370717	4.75715573054776	3.95951644275389
N	10.12036160155468	4.00202457645990	2.90086216587602
N	11.30718203202275	3.65115858397813	2.27710285313999
C	11.66324333026386	4.92454149742374	4.06158884863030
C	12.32216397601725	4.21218047417443	3.02107785016057
C	13.66704401467783	3.98589236202614	2.66292976059059
C	13.92789595871010	3.21178360062476	1.56271323614682
C	12.87488821802741	2.70994056020779	0.74102487167378
C	11.50769912457987	2.97565245704314	1.06524789282831
C	12.21077218825213	5.75804816594876	5.16748857652491
C	14.22509672584920	6.54148166993425	6.18924507135220
C	15.71304803524471	6.40882646985892	5.96918190600360
C	8.83069678990671	3.40117918551348	2.69528583857308
C	7.74217248685257	4.25269971569228	2.49234537382804
C	6.46793302301827	3.70005376026216	2.38339784153213
C	6.33172557331101	2.31090848951148	2.45103936357054
C	7.41992895553746	1.45947486463659	2.65647632274866
C	8.69006420020629	2.01458034254125	2.79559613561706
C	13.14917575618992	1.99630931360712	-0.45056194831463
O	1.09964719435694	0.34114628293783	4.01664666360885
O	2.88351701985351	-0.24241993470447	5.31470531274550
N	4.52870748331070	1.11279277099335	3.47441359546577
N	5.02084606522276	1.72181227506442	2.41083301247303
N	4.02349936883437	1.99150526232355	1.48708256762362
C	3.21971500931864	0.96435255674607	3.28427260824201
C	2.84503989423265	1.53568614854044	2.03594400688726
C	1.62913032961696	1.71968910542630	1.34613755157973
C	1.65788554259189	2.34982504310431	0.12951810642462

C	2.88952831459571	2.73640315026412	-0.47814687264652
C	4.13380759623609	2.50219795237686	0.18660857292510
C	2.40146647282316	0.28365675127956	4.32592036017257
C	0.17362673295586	-0.29891007356520	4.96688860435342
C	-1.23074249268890	-0.05950007823779	4.46560989077734
C	10.47599761337553	2.61480522203400	0.18365016144099
C	5.35431737696761	2.73393534819355	-0.46805625719103
C	10.79296085544913	1.93357624972853	-0.98301193857257
C	12.12638538243998	1.60382364431342	-1.29390593961697
C	2.92520006157194	3.29487314852606	-1.77867084807710
C	5.34116020494818	3.26347056919055	-1.75069129055643
C	4.12940850237908	3.56716825205979	-2.40128977953926
H	14.45170282937722	4.42346500835797	3.27444981921039
H	13.90919988020851	6.14314852175478	7.16337931804556
H	13.88025881393232	7.58129046990623	6.10282455389100
H	16.03706355074691	5.36194461532752	6.04846800677698
H	16.23766766214234	6.98806206955109	6.74206421881645
H	16.00879889836158	6.80382697660774	4.98750756400197
H	7.88402595408468	5.33191649310592	2.43712323633849
H	5.60131272857455	4.34777801105520	2.25181768568324
H	7.27632098190618	0.38130161447707	2.72499509381262
H	9.54958863227261	1.37120820472777	2.98388777034158
H	0.70886269559090	1.36694953562156	1.80417195230599
H	0.34644250685869	0.14592141087443	5.95639137632370
H	0.42638809950877	-1.36749312701622	5.00972939433261
H	-1.46263307834169	1.01359483043370	4.42524883109027
H	-1.93880798870290	-0.53774475908477	5.15704662081874
H	-1.37991592989917	-0.49753646900234	3.46904449890139
H	0.72998513865462	2.53363572913329	-0.41348486482317
H	9.44439447030553	2.88444700857716	0.38708913142405
H	14.19026855076504	1.78263360423659	-0.69727189689332
H	6.29987328451691	2.47603645185106	-0.00170103740651
H	9.99376350359578	1.65961050165023	-1.67276914830344
H	12.35075410419806	1.06024615522497	-2.21203578625048
H	1.97819606581287	3.48471757869264	-2.28582137790702
H	6.28802650098635	3.43547122941437	-2.26376799041916
H	4.14442354000913	3.99205775455785	-3.40505319653284
H	14.95804722219288	2.99634683160451	1.27629677419609

Table D.8. PBE0/TZVP-optimized Cartesian coordinates of TOQ dimer in water.

68
Coordinates from ORCA-job GeoOpt_PBE0

O	13.57316229077777	5.61844250012159	5.17452479266097
O	11.59495005546878	6.12118551444671	6.12187460457259
N	10.36053223176918	4.67891397658367	4.06668848417341
N	10.09643587994081	4.01320440615231	2.97774921078863
N	11.23650601122876	3.73466338794894	2.28500127855224
C	11.66214884364274	4.86006366637593	4.12217538264178
C	12.27146510453867	4.24757283890355	3.00390112469965
C	13.59505879102855	4.07571803364584	2.56732419467459
C	13.80477714407486	3.39302874142287	1.41323392128619
C	12.71933046050997	2.92913991379181	0.62124703091207
C	11.38156590136167	3.14202704366041	1.03195324966424
C	12.25684320182934	5.60646094989717	5.25668580566461
C	14.28878608615814	6.31928750940930	6.22509664847891
C	15.75867209865512	6.18296144069721	5.94412581688886
C	8.80761762316851	3.42350601768483	2.78294985270952
C	7.72987874678274	4.27161302630355	2.58359859362744
C	6.46309210739981	3.72313111976884	2.48015439888351
C	6.32870217035420	2.34508989464682	2.55377072468500
C	7.40596904519980	1.49711249377411	2.75592188260291
C	8.66934316628249	2.04789591770484	2.88604188520429
C	12.94295887277359	2.30598133776101	-0.61953737208585
O	1.08303289961906	0.46623426600544	4.02864900492583
O	2.77190414743954	0.10308145230373	5.47058645927425
N	4.49102979872414	1.25487239093284	3.58824921417848
N	5.02058672793224	1.76728526281616	2.51285014151499
N	4.08156744098490	1.94311268220928	1.54101137606269
C	3.20830036843350	1.07595057989697	3.35944414422620
C	2.89356403480236	1.52817797455500	2.05783638473053
C	1.71421974842152	1.63582027488366	1.30294040474486
C	1.79980943359247	2.15681479853474	0.05246945612651
C	3.05325854607942	2.51044869989391	-0.51731375471073
C	4.25093483247477	2.35637254252526	0.22096253458161
C	2.34434595419307	0.49176306836953	4.41331039479574
C	0.12247526697575	-0.08045604847622	4.97011727983968
C	-1.23738212090245	0.00679161289483	4.33551070775571
C	10.31558428492077	2.81440016834909	0.19197425508972
C	5.49364421955090	2.56415235025663	-0.38034104182904
C	10.57965549832078	2.22119175384527	-1.02206347160091
C	11.89178568469674	1.94741466118373	-1.42514067982206
C	3.14518226222424	2.95847725704459	-1.84731479929742

C	5.53901518871726	2.98575658311912	-1.69035165990307
C	4.36649350181046	3.20288122011629	-2.42341058673893
H	14.39775790862074	4.47941101665468	3.16557637155919
H	14.00962167272442	5.87195620923731	7.18052278100880
H	13.96305633487817	7.36090617346506	6.21730967262384
H	16.06521831906117	5.13526574905388	5.94859176431092
H	16.31731078721908	6.70551868140072	6.72315293453101
H	16.01979532795697	6.62670670569998	4.98151087206768
H	7.87401217628715	5.34313906986582	2.52371090138082
H	5.60054115446027	4.36387987012767	2.34495690688122
H	7.26059172358301	0.42638404841554	2.82493387927865
H	9.52634999779482	1.41056800407937	3.06600033835109
H	0.78218039258461	1.31447522299943	1.74262571044876
H	0.18402062543935	0.50114369486086	5.89108856903631
H	0.41106476049005	-1.11060939562249	5.18578929766419
H	-1.50624732625555	1.04263605375510	4.12031272060196
H	-1.97688437501476	-0.40155459784870	5.02746330035490
H	-1.27875263131186	-0.57202960255533	3.41107474072042
H	0.90691707665270	2.28034110684559	-0.54901798815112
H	9.29713883329103	3.04022347501632	0.46401550280870
H	13.96637507820022	2.13335657744869	-0.93172353713100
H	6.41266620157625	2.37621574906085	0.15118969950487
H	9.75420743528692	1.97257332756198	-1.67840517510679
H	12.07425760808671	1.47329653620884	-2.38162047915042
H	2.22895608222280	3.08688519240396	-2.41170634788455
H	6.50230024077446	3.14069227595475	-2.16166007949899
H	4.42743020379378	3.54283664258139	-3.44985479458686
H	14.81342104164064	3.21848615336755	1.05781004314615

Table D.9. BLYP/TZVP-optimized Cartesian coordinates of TOQ dimer in water.

68
Coordinates from ORCA-job GeoOpt_BLYP

O	13.56080493585887	5.64991129403225	5.25937033107008
O	11.49596867920983	6.19579145329767	6.08594316236916
N	10.36124738583423	4.69772306434413	3.98441051311553
N	10.13624745968623	3.99918054039625	2.87423146664031
N	11.34698866045386	3.67757034742767	2.24179723182196
C	11.67916794895370	4.86339876763224	4.10574441723832
C	12.35274381952417	4.20890645130000	3.02965775287677
C	13.70587673104057	4.01574886888831	2.67000892657008
C	13.98845335918953	3.30908874039182	1.52700932146201

C	12.94901733642308	2.83546206531655	0.66573983176564
C	11.56990769523150	3.06236823136823	0.99316690586941
C	12.21400702455172	5.64542307329580	5.26376128677355
C	14.23470874092329	6.39737012704452	6.35948877445930
C	15.73231649624559	6.26803092819645	6.14800491535414
C	8.84140834617670	3.38823744939675	2.64736460246706
C	7.74828665375773	4.23914386672811	2.44389088528015
C	6.47114923285786	3.68350683440678	2.33384913318982
C	6.33504414664324	2.29121734181414	2.39890987696981
C	7.42903906989958	1.43965338201010	2.59416625682722
C	8.70193095844257	1.99779455090961	2.73657774636285
C	13.24393002434277	2.18567606304856	-0.56213477848422
O	1.06285323177782	0.45689083847921	4.10608299357974
O	2.86762704411204	-0.11309056961535	5.39556164691872
N	4.51094173795449	1.14159271546377	3.47671939858703
N	5.01375984692013	1.69436703960024	2.37429832818829
N	3.99388951417607	1.94844918169276	1.44434304228348
C	3.19528248621412	1.00776853515904	3.30607832684058
C	2.81335728985062	1.53158903840328	2.03287571272491
C	1.58852780538531	1.69757698865756	1.34733813067171
C	1.60571240959802	2.26546559477441	0.09698747772686
C	2.83380910061893	2.61412674623937	-0.54840969218082
C	4.09001216883390	2.40386117185637	0.11404786952209
C	2.37666546745244	0.38403284834659	4.39232662959419
C	0.12320104675829	-0.12585966221445	5.10685623893875
C	-1.28863970176521	0.13303010767906	4.61174603087030
C	10.54868321711462	2.71866587270243	0.08713526759054
C	5.30586172897799	2.60934147822525	-0.56541088507872
C	10.88615473578835	2.10074954401414	-1.11203279774150
C	12.23158222808864	1.81502179737015	-1.43197999416763
C	2.85675068126998	3.11349989384739	-1.87777201857650
C	5.27901144994807	3.08081546105092	-1.87350001992478
C	4.05691178319304	3.35407278435974	-2.52673750871738
H	14.47896344213767	4.42533277491467	3.30973171193773
H	13.90605168255680	5.95551652738901	7.30673027281239
H	13.89263708178248	7.43708333030723	6.30928549639633
H	16.05164127817701	5.21919684784369	6.18632837894451
H	16.24799953589564	6.81196431290380	6.95043372159145
H	16.03967810906690	6.70321511137727	5.18890089195470
H	7.88603476947374	5.31635210033369	2.39276567294010
H	5.60636204530156	4.32963744706900	2.20597700550050
H	7.28987284993027	0.36324085800946	2.65621574622438
H	9.55978828940859	1.35553598047537	2.91996021408284
H	0.67198605472116	1.37967895383796	1.83058591671167

H	0.32433166214918	0.35704842620198	6.06935316551639
H	0.35040118710417	-1.19482820820243	5.18586375204286
H	-1.49056621046324	1.20800176072902	4.52825439475539
H	-1.99720626139899	-0.29446765868410	5.33353747390523
H	-1.46454639118416	-0.34289624193851	3.63900210369282
H	0.67291361150756	2.42803646708570	-0.43927114731277
H	9.51301141877355	2.95308836850718	0.29721710261846
H	14.28731054155136	2.00203499192966	-0.81237149569128
H	6.25446562733581	2.37634811266170	-0.09926628854034
H	10.09821483895572	1.84191375429431	-1.81651565707519
H	12.47122260295810	1.32292836158409	-2.37220859401475
H	1.90736465663325	3.28407281556993	-2.38181030165288
H	6.21807509526627	3.23105851099064	-2.40234464748389
H	4.06145980529884	3.73153166148577	-3.54693302477057
H	15.02244290154646	3.12645100798609	1.24117027726516

Table D.10. B3LYP/TZVP-optimized Cartesian coordinates of TOQ dimer in water.

68
Coordinates from ORCA-job GeoOpt_B3LYP

O	13.58759525656791	5.55506868613759	5.25142319390162
O	11.58525535021791	6.01900737044309	6.19138688004288
N	10.37816370623289	4.62570481493079	4.07836484383890
N	10.11828233715462	3.98976615913829	2.95877004381134
N	11.28087814417956	3.73329146379668	2.26372342021773
C	11.68203507988285	4.81124530168890	4.15034552245445
C	12.30740542586758	4.23521804886369	3.01505689797593
C	13.63907441705124	4.09207395476116	2.58410680603702
C	13.87161688371489	3.45444888156676	1.40456627952645
C	12.80042660568197	3.00020582311716	0.58272833955397
C	11.44963493641396	3.17766183600141	0.98871064814590
C	12.26232774150336	5.53100505040592	5.31617623216872
C	14.30130824931059	6.23923728891028	6.33621228348024
C	15.77995984509049	6.10971241112790	6.06239581697808
C	8.82265694657111	3.39854298907447	2.74421189044131
C	7.74497475715056	4.25242177932981	2.54339566094334
C	6.47222750627703	3.70868785238978	2.43695009751443
C	6.32877309439418	2.32790820919927	2.50734531160673
C	7.40674549613212	1.47366920798122	2.70441395135097
C	8.67588737468483	2.01976288489394	2.83931435120590
C	13.04582955715094	2.41803575779786	-0.67861870415422
O	1.04640399071219	0.54401973166728	4.08725677944389

O	2.76272359040959	0.17674162084284	5.51105910058562
N	4.47443951206749	1.26969382172961	3.57821235721945
N	5.00909190523569	1.75122487326689	2.47844297683873
N	4.04702792174938	1.92365565415560	1.50623035495429
C	3.18321482746683	1.10479898262374	3.36727077556638
C	2.85690462788148	1.53313311112785	2.05453962578625
C	1.66680970224598	1.63392591732555	1.31073354427828
C	1.73737783012706	2.11821293321294	0.04062416638934
C	2.98604485099562	2.44845716744547	-0.56026708517839
C	4.19882996431600	2.30652064192231	0.16745154339687
C	2.32289156931454	0.55453866504440	4.44970144970709
C	0.07578261826437	0.03073712428678	5.06182717060124
C	-1.29537477751991	0.13593677860494	4.43845757913954
C	10.39582962831005	2.84780480161497	0.12846351214183
C	5.43564975486578	2.50162961493692	-0.45811716194087
C	10.68188902972356	2.29585847639343	-1.10372209141841
C	12.00612322350340	2.06279435343346	-1.50550976507693
C	3.06189600133887	2.86430003864201	-1.90626275821517
C	5.46386933984302	2.89149375601377	-1.78196515498718
C	4.27845216849720	3.09137925917054	-2.50631283878407
H	14.43038139823011	4.48365619363123	3.20259050229191
H	14.01274144695847	5.76691237993614	7.27492909218299
H	13.97236569473135	7.27822553442852	6.34602641884380
H	16.08788002602791	5.06302308337100	6.04599853297246
H	16.33067009786301	6.61523869348873	6.85804218791832
H	16.04887268456480	6.57549806738654	5.11287533964659
H	7.89203385524356	5.32244737260848	2.48817581271795
H	5.61444226072866	4.35415255520591	2.30507949793714
H	7.25829205675224	0.40446992869000	2.76962747217569
H	9.52783704681213	1.37788417126485	3.01867775241155
H	0.73865829596876	1.33463787304290	1.77009140582290
H	0.16434560299660	0.63092323451799	5.96680167356358
H	0.34788898951289	-0.99942185276379	5.29065255120678
H	-1.54382128564209	1.17232594918552	4.20440827705417
H	-2.03440785394778	-0.24058247215561	5.14859131806854
H	-1.36161302403877	-0.46138135970253	3.52783554703187
H	0.83619797469644	2.23034101208774	-0.54836024922089
H	9.37182957100436	3.03866136754126	0.39961739500983
H	14.07330470196313	2.27284606027612	-0.98735570094261
H	6.36064255177450	2.32953191096016	0.06484034772557
H	9.86809580465104	2.04629343936205	-1.77205988565361
H	12.20551266881744	1.62192888178525	-2.47323381715227
H	2.13983872996411	2.98282340924498	-2.46082644532078
H	6.41937763799090	3.03603290505750	-2.26927454610789

H	4.32594623522295	3.40546643140776	-3.54058153039914
H	14.88586104257453	3.30877935712397	1.05584208272609

Table D.11. TPSS/TZVP-optimized Cartesian coordinates of TOQ dimer in water.

68
Coordinates from ORCA-job GeoOpt_TPSS

O	13.56432210053119	5.58687640279109	5.26103284877509
O	11.52974555168193	6.09531656995117	6.15783506503391
N	10.35485663793313	4.67001934090090	4.03618038852944
N	10.11654756077181	4.01248126889107	2.91254471376408
N	11.30113509683158	3.73101399272062	2.24236161602832
C	11.67033545077643	4.84835634840262	4.12712598318046
C	12.31944794701274	4.24851750478025	3.01302793788110
C	13.65968926640401	4.09324707479389	2.60207822002032
C	13.91167355380679	3.43564487497126	1.42901250308981
C	12.85282867338779	2.97020272394444	0.59361544654077
C	11.48884196659895	3.15408795008919	0.97365505451175
C	12.22552533503614	5.58243634949816	5.29838303183321
C	14.25414517862714	6.27761119785132	6.37781970775531
C	15.74124808217286	6.13764715463608	6.13390258522013
C	8.82444168380336	3.40280334907818	2.68822023442691
C	7.73902654653178	4.25193961439760	2.47602463178194
C	6.46743020057142	3.69497496637265	2.36929592982235
C	6.33994058233273	2.30819774980913	2.45455477045433
C	7.42636835142353	1.45816800073753	2.65814004913604
C	8.69462231503951	2.01755857909940	2.79020360492753
C	13.11686614102158	2.36761942131147	-0.65924808624034
O	1.06400505755145	0.51147856538389	4.10387859715933
O	2.82252528991632	0.04798821562049	5.47977264798528
N	4.50509276911611	1.20895470439230	3.54619048618474
N	5.02090493298447	1.71549823670060	2.43668093273904
N	4.03295464843105	1.92070538536579	1.48103254904113
C	3.19822121999191	1.05637068305581	3.34927436377800
C	2.84514526897988	1.51822838670021	2.05101299776583
C	1.64218071474499	1.63942025059985	1.32469475334428
C	1.69328994069450	2.14849475688618	0.05540418510289
C	2.93431534686164	2.48405741266670	-0.56365460194886
C	4.16592453423589	2.32553147176948	0.14104894325351
C	2.36052116321831	0.47946608409840	4.43756717946382
C	0.11210323069174	-0.02895790506442	5.10560857445425
C	-1.28082933848959	0.17933394581871	4.55082096213182

C	10.44351860138743	2.81036072401167	0.10364388177609
C	5.39746473995153	2.52560970586870	-0.50024390009906
C	10.75058691488510	2.23805772362808	-1.12101014038363
C	12.08491407553022	1.99959001850209	-1.49982427578826
C	2.99090986080175	2.92234806439452	-1.90812327025834
C	5.40410855465006	2.93770495696023	-1.82396038256574
C	4.20383524630104	3.15441171857634	-2.52670787743786
H	14.44487267998493	4.49384147423682	3.22845411534661
H	13.93329381627894	5.79763599483168	7.30524605862577
H	13.92268356136579	7.31857349579979	6.37148127203173
H	16.04082661440640	5.08544436646210	6.12081873155956
H	16.27700347115977	6.63854463211840	6.94685658428470
H	16.03241984816575	6.60748257861908	5.18972561840901
H	7.87916670321101	5.32557281277120	2.41469237964307
H	5.60146092840984	4.33299870069966	2.23072848229614
H	7.28471187229055	0.38559198832757	2.73163938752497
H	9.55500989384688	1.38355165149822	2.97537698939668
H	0.71774554263265	1.33261813305600	1.79464696353940
H	0.27759565382921	0.51288584101571	6.03934772367453
H	0.35558487273942	-1.08456453171211	5.24743877264212
H	-1.48889527539812	1.24264929164862	4.39929095051661
H	-2.00538867942981	-0.21561175672054	5.27057483472801
H	-1.41197770256669	-0.35278339681275	3.60404214367217
H	0.78007066926351	2.27477326104341	-0.51771569736291
H	9.41263290019653	3.00446658521300	0.36212397587290
H	14.15222566745283	2.21813814375418	-0.95058485887795
H	6.33187449068054	2.34060099302032	0.00981813227226
H	9.94436058886843	1.97727837873062	-1.79944942951285
H	12.29980447943585	1.54373015986091	-2.46086828370711
H	2.05722937161185	3.05284103559359	-2.44663553648815
H	6.35503457576722	3.08593645993715	-2.32616832791826
H	4.23578684887823	3.48495612434055	-3.55971941389778
H	14.93424581418624	3.28372126370320	1.09807046955798

Table D.12. PBE/TZVP-optimized Cartesian coordinates of TOP monomer in water.

40
Coordinates from ORCA-job GeoOpt_PBE

O	13.63973271549595	5.46680891666839	5.32088673936176
O	11.71396736824273	5.35610609352554	6.55043157831788
N	10.46256088644604	4.18683406913550	4.30906417604563
N	10.19081541908225	3.63979891994140	3.14409382281117
N	11.32330615517196	3.59996848229973	2.33333399070202
C	11.75472883895036	4.51969974982253	4.31616717220280
C	12.34132595225627	4.17001579140725	3.07092155572647
C	13.63519191762449	4.29294628173485	2.51557838815671
C	13.84764761479957	3.83310338512938	1.24516216744022
C	12.80305637527116	3.21754595916851	0.48941177662060
C	11.48949449086247	3.07105375481984	1.04010794469422
C	12.34218203431257	5.15860574646871	5.53566424510055
C	14.32756850292592	6.11001770940417	6.45495370937959
C	15.74632152356298	6.40384164010396	6.02896458994993
C	8.84499546888706	3.26115785453419	2.78533503456240
C	8.05486492498550	4.17053059771297	2.07710115854210
C	6.73960072196813	3.81273608676599	1.77634411710369
C	6.23554065159365	2.57511569862538	2.19142132707234
C	7.03789821091915	1.68900358751583	2.91766522525785
C	8.35711887589356	2.02872442567072	3.22486830582611
C	13.03870757276737	2.71625631894488	-0.81260503814637
C	10.47595318325200	2.42405921072044	0.31284811187079
C	10.75737610445425	1.94425182676866	-0.95897403045357
C	12.03448392262859	2.09238453003147	-1.52932492822569
H	14.41073343223076	4.75243818539676	3.12375713866980
H	14.27671358815447	5.42219034125846	7.31067545205354
H	13.76792820415352	7.01962634523054	6.71462466477621
H	16.28960737381323	5.48412531265762	5.77124664555713
H	16.27760667800746	6.88407601278472	6.86319298646255
H	15.77477925644562	7.09106274686458	5.17188397143428
H	8.45726169596006	5.13841090795635	1.77580302400715
H	6.10309778950101	4.50998022792757	1.22992827908192
H	6.63562274684757	0.73345061232225	3.25538058224561
H	8.99503567813999	1.35677524341668	3.80075451356289
H	9.48242781032069	2.28958670939390	0.72813781412357
H	14.03771738098427	2.82868624216313	-1.23769514288638
H	9.96757249566698	1.44150101113511	-1.51847466505569
H	12.23191198825808	1.71076249161808	-2.53127148437889
H	14.83444333644240	3.92119970897763	0.78782023656264
H	5.20485273571984	2.30500484697633	1.95785718786450

Table D.13. PBE0/TZVP-optimized Cartesian coordinates of TOP monomer in water.

40
Coordinates from ORCA-job GeoOpt_PBE0

O	13.60534033931538	5.47902193796383	5.30928840198154
O	11.67701765855595	5.49812142593968	6.46730562359349
N	10.45462036563234	4.27341989085873	4.26914528505868
N	10.19375923368861	3.71502606177973	3.12666545323265
N	11.31295590119414	3.63620594770068	2.35084363027177
C	11.73672058227082	4.57917029059583	4.28908496086977
C	12.32830420514256	4.18926691056336	3.07043104526818
C	13.62594452313593	4.26813396000265	2.53627378550967
C	13.83486693204820	3.78031479689416	1.28952802743325
C	12.78038184064555	3.19190432653703	0.53896661671426
C	11.47058043293675	3.09892781385035	1.07081710498225
C	12.31952103843221	5.23442191777074	5.48194650606905
C	14.30072981683191	6.12456705024093	6.40703845609204
C	15.73458334412785	6.30623049284256	5.99453461360226
C	8.86457386830010	3.29789876357887	2.78600448787106
C	8.03744304142262	4.19205148365843	2.12392145084877
C	6.74416215076066	3.79073011113968	1.82345991585048
C	6.30401111951606	2.52447078023333	2.19220119317923
C	7.14708494596902	1.65226503045939	2.87118722558955
C	8.44419222242816	2.03680934778328	3.17838616017624
C	13.01820474245177	2.67775024030207	-0.74770613265744
C	10.44730969645838	2.49728969344045	0.33369698205478
C	10.72504225279819	2.00511982679543	-0.92229942665427
C	12.00842407547157	2.09456598313145	-1.47095147338579
H	14.40764769310799	4.71397974968472	3.13232670954482
H	14.20408941976946	5.48892330909875	7.28875931851147
H	13.80686114886170	7.07712002161283	6.60670735039551
H	16.21101906123222	5.34553486264789	5.78978746460061
H	16.27807343261540	6.78974458141546	6.80918170338095
H	15.81246586699216	6.93915033691912	5.10837827989700
H	8.40067510987311	5.17398968499896	1.84668953871243
H	6.07926915431947	4.47106592142983	1.30466089650605
H	6.79620907009314	0.67044538108996	3.16506945543863
H	9.11879929016744	1.37283653769729	3.70568639756436
H	9.44708398384271	2.40503731131148	0.72273726874574
H	14.02091239368817	2.75186345258295	-1.15289621244271
H	9.92695280022744	1.53961357460349	-1.48803627895364
H	12.20246388793704	1.70202039601662	-2.46157908380970
H	14.82090371510936	3.82746795618467	0.84247299278231
H	5.29255126562839	2.21696642164332	1.95332664957468

Table D.14. BLYP/TZVP-optimized Cartesian coordinates of TOP monomer in water.

40
Coordinates from ORCA-job GeoOpt_BLYP

O	13.63410418302318	5.49801922050675	5.32942871357782
O	11.66006868792592	5.53090436616263	6.48861472470644
N	10.43611713998560	4.28867971854732	4.27204178728375
N	10.16190352374610	3.71464120866995	3.11140652793811
N	11.31882967183773	3.62490149288073	2.31822151762969
C	11.73908282869327	4.58975065119984	4.28962768351511
C	12.34222042221405	4.18558575274526	3.06362666554389
C	13.65272940573785	4.25033477671107	2.53444560968212
C	13.88025262531214	3.75097401316312	1.27839226500380
C	12.82771600496909	3.16764234239976	0.50392169613840
C	11.49063442961070	3.08641790300988	1.02327709630925
C	12.31824255467236	5.25433876973718	5.49521848425411
C	14.34046359645036	6.16437975890471	6.45942098013178
C	15.79241627586582	6.33994592650472	6.05031244895722
C	8.81323548159516	3.28673648182733	2.77859103089344
C	7.95321976357339	4.20068342477490	2.15740838469803
C	6.63936610567968	3.79858144440624	1.88711902470969
C	6.20934142611181	2.51095576890041	2.24232640694559
C	7.08567271755922	1.61704069952897	2.87615949389452
C	8.40317884474764	2.00168667298756	3.15451120524687
C	13.07430281775221	2.64968937779546	-0.79461344505603
C	10.46110649823528	2.50155212197490	0.25959305232597
C	10.75084091168309	2.00729830738582	-1.00789107146252
C	12.05542770118276	2.08007100042042	-1.54082175860399
H	14.43596620733514	4.69393468128049	3.13815379669781
H	14.22943774534506	5.52449287886495	7.34176074949767
H	13.83995386264982	7.12153523195681	6.64409788487971
H	16.27166612230145	5.37259840586591	5.85461199151001
H	16.33206132192242	6.82820007535825	6.87306832650386
H	15.88123448207139	6.97325848215176	5.15874364560344
H	8.30345843384185	5.19563670032243	1.89296177493079
H	5.95365639445849	4.49350443826681	1.40575864763072
H	6.74616317454446	0.62259475978350	3.15897371687634
H	9.09627590675778	1.32357249124082	3.64786952080479
H	9.45062574541489	2.42550224290611	0.63649618347017
H	14.08660656495726	2.71150369496934	-1.19073236235979
H	9.95001071603215	1.55750450234537	-1.59133641501378
H	12.25858051729909	1.68845626121219	-2.53535377026126
H	14.87916248723760	3.79126468562126	0.84831463970839
H	5.18641832266674	2.20507284970909	2.02931548925808

Table D.15. B3LYP/TZVP-optimized Cartesian coordinates of TOQ monomer in water.

40
Coordinates from ORCA-job GeoOpt_B3LYP

O	13.60958452167168	5.48669123118645	5.31824440178807
O	11.66400967639240	5.51529103624033	6.46666482967069
N	10.44631477641777	4.28204332785295	4.26200613878209
N	10.17948284913470	3.71611955538082	3.11368864393952
N	11.31551191320795	3.63159589026839	2.33507911250165
C	11.73325094896619	4.58428389815049	4.28472094595688
C	12.33136006149410	4.18620251773893	3.06570758983591
C	13.63537684068882	4.25689105502744	2.53846678248184
C	13.85552994263137	3.76403468088257	1.29153514327491
C	12.80420980750738	3.18038022825305	0.52693277393217
C	11.48323189143515	3.09531958298160	1.04790125419156
C	12.31379139957069	5.24438109877769	5.48243584262390
C	14.31504155691696	6.14140515363570	6.42536439169204
C	15.75648645750614	6.31311129606426	6.01034537384456
C	8.84054143989257	3.29206895935790	2.78046726076788
C	7.99365602156799	4.19504142509950	2.14849173383702
C	6.69140714447332	3.79441130784239	1.86926946118937
C	6.26067339369965	2.51834710163561	2.22823245345449
C	7.12377261570542	1.63540358572070	2.87462761960086
C	8.42948363695346	2.02007012184768	3.16082709286522
C	13.04930841543152	2.66622674467598	-0.76317818315696
C	10.45967679804833	2.50639981076661	0.29444433439981
C	10.74507747054479	2.01526605617683	-0.96399193450446
C	12.03861146712477	2.09402676542156	-1.49984375763523
H	14.41411131020050	4.69953888273748	3.13815299447075
H	14.21240938030681	5.50860422826351	7.30660343356991
H	13.82445442926888	7.09589793479277	6.61553389852400
H	16.22782519404013	5.34911761640331	5.81183071815244
H	16.30118937075358	6.79827392648781	6.82304156304132
H	15.83873861530350	6.94046611585449	5.12118507075960
H	8.34596676097167	5.18203368077683	1.88024530559221
H	6.01454114754488	4.48091874147720	1.37701992730105
H	6.78253309049608	0.64882877947874	3.15996391713888
H	9.11444556146925	1.34968100401708	3.66337072006333
H	9.45635746770154	2.42428710947850	0.67200244400058
H	14.05461520497042	2.73261030487266	-1.15995989261353
H	9.94899327406152	1.56180334512310	-1.54002192350633
H	12.23856563078061	1.70383292550494	-2.48907280712168
H	14.84597006483428	3.80604634911431	0.85684384275203
H	5.24564407331316	2.21249020763184	2.00786382654163

Table D.16. TPSS/TZVP-optimized Cartesian coordinates of TOQ monomer in water.

40
Coordinates from ORCA-job GeoOpt_TPSS

O	13.63092635565890	5.46612703868414	5.33038545457564
O	11.67849758380053	5.48994873742865	6.50812507976344
N	10.44781490192737	4.24324401005596	4.30084927805738
N	10.17994703811758	3.67918062687164	3.14070657400465
N	11.31872132134299	3.62027820042627	2.34235486830890
C	11.74257974274195	4.56472673423148	4.30821213759224
C	12.33684294855523	4.19223512630683	3.07505694965047
C	13.63154711719335	4.30033617503835	2.52309659314277
C	13.84371434456773	3.82893854142221	1.25852846361073
C	12.79730916420100	3.21527164554358	0.50675024949803
C	11.48365993853835	3.07953607682378	1.05151044985888
C	12.32447414013471	5.22106395249442	5.50920890185146
C	14.33991274379154	6.12475886630383	6.45329229626046
C	15.77537435228472	6.32508729370109	6.01472859059913
C	8.83139555574339	3.27833582875517	2.78632556230860
C	8.03080867614483	4.17864828558559	2.08277409908622
C	6.72211977987057	3.79847757554836	1.78337771062205
C	6.24135477656159	2.55058715727452	2.19297563281233
C	7.06047819301356	1.67347089166511	2.91054269763315
C	8.37352235673640	2.03376820179869	3.21779710588700
C	13.03629153144163	2.70675695653506	-0.79144115999021
C	10.46831136754535	2.43626259233023	0.32644062107252
C	10.75069651110690	1.95021289829637	-0.94099826605363
C	12.03080688577322	2.08725244946264	-1.50750611944927
H	14.41149454776316	4.75641342366577	3.11719278926116
H	14.25462176017925	5.46896798067286	7.32282522180794
H	13.82775157517416	7.06820961285506	6.65606948948698
H	16.25791512147518	5.36811922375412	5.79426427405679
H	16.32517082638008	6.80802305161317	6.82969314488153
H	15.83080357500061	6.96900310100799	5.13175555387691
H	8.42142922536664	5.14374575625374	1.77848985204968
H	6.07856060293749	4.48127096844674	1.23697432807507
H	6.68109732918229	0.70931350482266	3.23376462901072
H	9.02774854571137	1.36852685766086	3.77211581348281
H	9.47729776118302	2.30710164548024	0.73465226138774
H	14.03376595235803	2.81183329733731	-1.20823724316191
H	9.96232033170510	1.45317553629265	-1.49744088268117
H	12.22654966496744	1.70123117038893	-2.50268895199439
H	14.82667239051293	3.90713829963862	0.80434020788578
H	5.22144508630979	2.26286429052530	1.95617808587143

Table D.17. PBE/TZVP-optimized Cartesian coordinates of TOP monomer in the gas phase.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.54142636593925	6.35242066175758	4.63257439677545
O	11.47142530604380	7.15409913707372	5.19452976994560
N	10.36528860562670	5.01439686834600	3.69998276825322
N	10.18143821017400	3.95193519679944	2.93890693583918
N	11.39248428175498	3.43907158530724	2.51250789213602
C	11.67948431548795	5.21706945687012	3.81375410894011
C	12.38558668219312	4.22308059379384	3.07512754204083
C	13.72566683500424	3.89615025842395	2.79962898964463
C	13.98891809284532	2.82253005272914	1.96934841177507
C	12.93102711458707	2.07205077843481	1.39979057209489
C	11.62025126829401	2.38050640427666	1.67383042602313
C	12.19054499831052	6.36059530964781	4.63323667420911
C	14.17926479678222	7.43398106223168	5.40600030003583
C	15.67513042117468	7.26082367596038	5.29478941278286
C	8.90268353534409	3.44536086096408	2.52690284288942
C	8.05304546420756	4.29479278176915	1.81235862796614
C	6.79722221773898	3.81517723225793	1.43851803191601
C	6.41423591687222	2.51097124609262	1.76820588633584
C	7.27720879519713	1.67812443273945	2.48849638929343
C	8.53255141189863	2.14322924206372	2.88299236226898
H	14.50846376970760	4.50537834529472	3.24709515382322
H	15.01977460351859	2.55117381593884	1.74092531380582
H	13.13906359868702	1.23593927861730	0.73257825569649
H	10.74800181315082	1.86479041824698	1.27800869934452
H	13.82082908800549	7.36053005618574	6.44251759939615
H	13.83284967536538	8.39070759148416	4.99038328961338
H	16.00230735628683	6.29841127567088	5.71190953735927
H	16.17102524843999	8.05874450469012	5.86515202950856
H	16.01298993903039	7.33582589446596	4.25195090448079
H	8.36704478723267	5.31055113734378	1.57071888268993
H	6.11588015794108	4.46475798386832	0.88751225635408
H	5.43102643647441	2.14415362724461	1.47103616442255
H	6.96586329235654	0.67004339178858	2.76368317668012
H	9.19564164632668	1.51780985762067	3.48376840565934

Table D.18. PBE/TZVP/RIJCOSX-optimized Cartesian coordinates of TOP monomer in the gas phase.

34
Coordinates from ORCA-job RIJ

O	13.54142636593925	6.35242066175758	4.63257439677545
O	11.47142530604380	7.15409913707372	5.19452976994560
N	10.36528860562670	5.01439686834600	3.69998276825322
N	10.18143821017400	3.95193519679944	2.93890693583918
N	11.39248428175498	3.43907158530724	2.51250789213602
C	11.67948431548795	5.21706945687012	3.81375410894011
C	12.38558668219312	4.22308059379384	3.07512754204083
C	13.72566683500424	3.89615025842395	2.79962898964463
C	13.98891809284532	2.82253005272914	1.96934841177507
C	12.93102711458707	2.07205077843481	1.39979057209489
C	11.62025126829401	2.38050640427666	1.67383042602313
C	12.19054499831052	6.36059530964781	4.63323667420911
C	14.17926479678222	7.43398106223168	5.40600030003583
C	15.67513042117468	7.26082367596038	5.29478941278286
C	8.90268353534409	3.44536086096408	2.52690284288942
C	8.05304546420756	4.29479278176915	1.81235862796614
C	6.79722221773898	3.81517723225793	1.43851803191601
C	6.41423591687222	2.51097124609262	1.76820588633584
C	7.27720879519713	1.67812443273945	2.48849638929343
C	8.53255141189863	2.14322924206372	2.88299236226898
H	14.50846376970760	4.50537834529472	3.24709515382322
H	15.01977460351859	2.55117381593884	1.74092531380582
H	13.13906359868702	1.23593927861730	0.73257825569649
H	10.74800181315082	1.86479041824698	1.27800869934452
H	13.82082908800549	7.36053005618574	6.44251759939615
H	13.83284967536538	8.39070759148416	4.99038328961338
H	16.00230735628683	6.29841127567088	5.71190953735927
H	16.17102524843999	8.05874450469012	5.86515202950856
H	16.01298993903039	7.33582589446596	4.25195090448079
H	8.36704478723267	5.31055113734378	1.57071888268993
H	6.11588015794108	4.46475798386832	0.88751225635408
H	5.43102643647441	2.14415362724461	1.47103616442255
H	6.96586329235654	0.67004339178858	2.76368317668012
H	9.19564164632668	1.51780985762067	3.48376840565934

Table D.19. PBE/TZVP-optimized Cartesian coordinates of TOP monomer in the gas phase in the excited state.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.56098144490194	6.39682703436964	4.66303273431844
O	11.54491551491302	7.04384201590139	5.19433098287976
N	10.43381305598020	4.99807967008734	3.65070419235453
N	10.22721762924740	3.96511023899214	2.88893014331604
N	11.42939325820685	3.41908827176763	2.43670392323427
C	11.77119651120490	5.14668432643337	3.73284156307395
C	12.47491752078632	4.14344747828081	2.97745828376844
C	13.78431345964747	3.78451936854736	2.68958624811714
C	13.98272162554316	2.65280850554617	1.81559331933569
C	12.90985864880071	1.97281074327311	1.28456835275910
C	11.57588438262271	2.33091409484608	1.57481288577481
C	12.28522483662362	6.21689265949322	4.54667105785794
C	14.11582326278778	7.51744825952800	5.52417911056041
C	15.61180651999580	7.42060793844039	5.44322014725059
C	8.92919439024504	3.47158968171214	2.52179707770203
C	8.05842399639974	4.32928744556466	1.84456471177847
C	6.78378365698831	3.86379733501506	1.51972723851682
C	6.40339105797948	2.56122044778178	1.85749607062119
C	7.29024612920286	1.71713261516213	2.53371111479508
C	8.56485760895285	2.16889647216788	2.87855618726107
H	14.61689036380745	4.34432748320013	3.09707258262992
H	14.99721170255635	2.34361471794890	1.57054011138080
H	13.06606766804705	1.12781343695660	0.61432076730528
H	10.67663905730320	1.88569240750827	1.17080231758840
H	13.72204516994223	7.35291010865519	6.53609423402966
H	13.71661053711799	8.45185706197927	5.10826213169142
H	15.97852503365218	6.46672168347242	5.84278000436472
H	16.03790846921070	8.22810814803748	6.05655007024607
H	15.97094152633159	7.55599273188459	4.41526236510277
H	8.37545462105787	5.33925725354220	1.58417148919961
H	6.08822168090101	4.52048779366538	0.99558086108856
H	5.40707466690607	2.20282348525035	1.59582632438305
H	6.98609810277737	0.70642756302155	2.80737546609334
H	9.25799293735875	1.52814553796737	3.42559793962061

Table D.20. PBE/TZVP/RIJCOSX-optimized Cartesian coordinates of TOP monomer in the gas phase in the excited state.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.56098144490194	6.39682703436964	4.66303273431844
O	11.54491551491302	7.04384201590139	5.19433098287976
N	10.43381305598020	4.99807967008734	3.65070419235453
N	10.22721762924740	3.96511023899214	2.88893014331604
N	11.42939325820685	3.41908827176763	2.43670392323427
C	11.77119651120490	5.14668432643337	3.73284156307395
C	12.47491752078632	4.14344747828081	2.97745828376844
C	13.78431345964747	3.78451936854736	2.68958624811714
C	13.98272162554316	2.65280850554617	1.81559331933569
C	12.90985864880071	1.97281074327311	1.28456835275910
C	11.57588438262271	2.33091409484608	1.57481288577481
C	12.28522483662362	6.21689265949322	4.54667105785794
C	14.11582326278778	7.51744825952800	5.52417911056041
C	15.61180651999580	7.42060793844039	5.44322014725059
C	8.92919439024504	3.47158968171214	2.52179707770203
C	8.05842399639974	4.32928744556466	1.84456471177847
C	6.78378365698831	3.86379733501506	1.51972723851682
C	6.40339105797948	2.56122044778178	1.85749607062119
C	7.29024612920286	1.71713261516213	2.53371111479508
C	8.56485760895285	2.16889647216788	2.87855618726107
H	14.61689036380745	4.34432748320013	3.09707258262992
H	14.99721170255635	2.34361471794890	1.57054011138080
H	13.06606766804705	1.12781343695660	0.61432076730528
H	10.67663905730320	1.88569240750827	1.17080231758840
H	13.72204516994223	7.35291010865519	6.53609423402966
H	13.71661053711799	8.45185706197927	5.10826213169142
H	15.97852503365218	6.46672168347242	5.84278000436472
H	16.03790846921070	8.22810814803748	6.05655007024607
H	15.97094152633159	7.55599273188459	4.41526236510277
H	8.37545462105787	5.33925725354220	1.58417148919961
H	6.08822168090101	4.52048779366538	0.99558086108856
H	5.40707466690607	2.20282348525035	1.59582632438305
H	6.98609810277737	0.70642756302155	2.80737546609334
H	9.25799293735875	1.52814553796737	3.42559793962061

Table D.21. PBE0/TZVP-optimized Cartesian coordinates of TOP monomer in gas phase.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.51857767939671	6.33906146725579	4.63598162523545
O	11.48057660893882	7.11987622888837	5.21226582419190
N	10.37475273252727	5.01075788769721	3.70488204028978
N	10.19458754464623	3.96991946345016	2.93997585485318
N	11.38441732015618	3.47218806525905	2.51452786023101
C	11.67180729803094	5.21625260250085	3.81838918568779
C	12.37174223794715	4.23813379345746	3.06946464183640
C	13.70400396409264	3.91780923849098	2.78469679655518
C	13.95574753507167	2.85903574501937	1.95182083554401
C	12.89697819229689	2.11659709221053	1.38971988929670
C	11.60115076564755	2.42462983453384	1.67393027758405
C	12.18848883551148	6.34391296164175	4.64356167933154
C	14.15910823408433	7.39290823620426	5.40475938810327
C	15.64622623652496	7.21786903628835	5.27892687322376
C	8.92380904033481	3.45671591667287	2.53551690777885
C	8.10498026480360	4.26206472585912	1.75824448508091
C	6.86193361433925	3.77400037901899	1.38690472958827
C	6.46402022918983	2.50360949300192	1.78287403998320
C	7.29802022671882	1.71348248562797	2.56441790168011
C	8.53904976743845	2.19028075303618	2.95617848971624
H	14.48511965156682	4.51874578441180	3.22821154078250
H	14.97706939318622	2.58901158192076	1.71366572911904
H	13.09774085373542	1.28890632515078	0.72184035726837
H	10.72963146265875	1.91607562981242	1.28702318271862
H	13.81709198745229	7.31145218500540	6.43823157916369
H	13.81883208534424	8.35184678492140	5.00962355106278
H	15.96898043380980	6.25399904677790	5.67679778346901
H	16.14818520501314	8.00047162256411	5.85066958759647
H	15.97048221254534	7.30289525294744	4.24007379570443
H	8.43044744108827	5.25287109295647	1.46576679043009
H	6.20227722436853	4.38972114945533	0.78742060364372
H	5.49107082417292	2.12929107064349	1.48799521487861
H	6.97492499764213	0.73141787793973	2.88688603876941
H	9.18381394771852	1.59937320537796	3.59747692960167

Table D.22. PBE0/TZVP/RIJCOSX-optimized Cartesian coordinates of TOP monomer in gas phase in the excited state.

34
Coordinates from ORCA-job PBE_Ps1_RIJ

O	13.27038363317460	6.26475489285802	4.39359748872387
O	11.28420430621429	7.32264955986657	4.26204757723713
N	10.50518093821099	5.44998725416061	2.34660680321175
N	10.36143853251446	4.21679449052081	1.60326739039148
N	11.52846582221574	3.44676276848361	1.85802880857004
C	11.63204050668682	5.33080891619487	3.03132071678652
C	12.31945056168880	4.13581438247254	2.76066529450861
C	13.54823147832982	3.55116943985757	3.10596151541811
C	13.95890346519021	2.43161823100364	2.43327311587944
C	13.18216753664904	1.88189495125886	1.37607266375868
C	11.97200057023715	2.43188073094981	1.08127932759269
C	12.01805070129481	6.42498692750639	3.96177961064440
C	13.75800757713338	7.28407825141336	5.29946765113476
C	15.20215557782441	6.98876473164354	5.59540606770292
C	9.16480597260830	3.65423538827852	2.05955030871307
C	8.02591335017794	4.54688827302415	2.04781507513366
C	6.78958700894107	4.09878941270541	2.44219651743808
C	6.63918955441838	2.77583625488498	2.83412877223436
C	7.76399815645646	1.89646599876796	2.84296752791755
C	9.01307280942999	2.32892086451983	2.46608349420977
H	14.15970209911616	4.03243861581031	3.85573458863752
H	14.91268508828843	1.98089821313978	2.67705347914501
H	13.54456957832178	1.04814788891169	0.79173985754400
H	11.32245346800329	2.10221606140127	0.28181630208756
H	13.13799277047855	7.26576250674157	6.19809699663236
H	13.62676502037876	8.25640414832344	4.82157840561063
H	15.32137087090856	6.01465330145209	6.07390044483553
H	15.59219922602461	7.74686391750294	6.27738175084737
H	15.80491163949958	7.01241653772702	4.68597265307944
H	8.18828460468318	5.56366510478511	1.72061258698090
H	5.94104496426010	4.77060686210039	2.43251188330528
H	5.67006359386281	2.39554375610509	3.13373644203727
H	7.62109959317165	0.86895601403235	3.15538587596608
H	9.85525547160587	1.65350936759589	2.49168501608419

Table D.23. BLYP/TZVP-optimized Cartesian coordinates of TOP monomer in gas phase.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.55223161547800	6.35634010026435	4.65213427659056
O	11.46903853266558	7.12750037450956	5.24242803267475
N	10.36509505288294	4.99919260143168	3.71980752785824
N	10.17243255453201	3.94168861419632	2.93679002349942
N	11.39787726491837	3.44168134534606	2.48678935802138
C	11.68234691315805	5.20917854844913	3.82315665167591
C	12.39166605605324	4.23054065874394	3.05675257084053
C	13.73547427772699	3.91776614983229	2.76366244875334
C	14.00367744886015	2.85519456494162	1.91551015176631
C	12.94693563517732	2.09679166014739	1.34447939697102
C	11.63233132711883	2.38927194385181	1.63201355810310
C	12.19315465820772	6.34928350496686	4.66138838882370
C	14.20178581717794	7.44158284496646	5.44575765613479
C	15.70474957899517	7.26458820249481	5.32636506423369
C	8.88126189938241	3.43178729977627	2.52665205901274
C	8.06193194935409	4.24751662374140	1.73447728966495
C	6.80017503166163	3.76760018029364	1.36696627953880
C	6.38007699766540	2.49587275414738	1.78187043043915
C	7.21284139054014	1.69593781330249	2.57767094410603
C	8.47400318933417	2.16262275972236	2.96450556427075
H	14.51617106117935	4.52449506087027	3.21064161677412
H	15.03319192078997	2.59817867970254	1.67618698078088
H	13.15955361250608	1.27120602957273	0.66946469937464
H	10.76616612653073	1.86709402743264	1.23983643891924
H	13.84537783253978	7.35213858089280	6.47835857111736
H	13.85477164213944	8.39792029889310	5.03786877740972
H	16.02993881534366	6.29848961316454	5.73211741335194
H	16.20256805009756	8.05535620194778	5.90258385336727
H	16.03811906880869	7.34855845431428	4.28448769719789
H	8.39943750871574	5.23538801973431	1.42891497006125
H	6.14559000553355	4.38950657114883	0.75910234112215
H	5.39653703501526	2.13091296634245	1.49218074880809
H	6.87564467379008	0.71776237771732	2.91410984981597
H	9.11349150411990	1.56623858914057	3.61369037892033

Table D.24. BLYP/TZVP-optimized Cartesian coordinates of TOP monomer in gas phase in the excited state.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.52838493608481	6.39356048990032	4.64987766649713
O	11.55001412515265	6.85080928756978	5.45494123754497
N	10.40081605814568	4.69209101095070	4.07814100198186
N	10.18531230415320	3.51408394812347	3.41287395808750
N	11.43740699258484	3.15671537259754	2.70645482353358
C	11.74597716611412	4.93593958369673	3.99262343379707
C	12.43477537407632	4.03445049082571	3.16904305060947
C	13.75306477123321	3.81496904714937	2.66888132787905
C	13.97868648010251	2.77040151367065	1.79924047924895
C	12.91798694832148	1.90459761385464	1.37613486785849
C	11.64038755926470	2.12314738921535	1.85253366372829
C	12.26808637955797	6.07947061368078	4.71579107440871
C	14.10152369257506	7.59966473523858	5.42534300507235
C	15.57963832462312	7.63819084388340	5.12505871916680
C	8.91775978642358	3.28597293596872	2.74663745114569
C	8.29071151689530	4.31648765564671	2.03002603964092
C	7.05584321669755	4.06557982199223	1.42414224103953
C	6.46211045508089	2.79897114529022	1.52807348037134
C	7.09667629569929	1.77858609283592	2.24891441667833
C	8.32732214434944	2.02001139077614	2.87088838561205
H	14.56095421191262	4.46775900151258	2.98078572928958
H	14.98390142885775	2.59918249288332	1.42042749253324
H	13.09966180261362	1.08320104418298	0.69058577088531
H	10.77057144637647	1.53048025344291	1.58586869152042
H	13.87240471397241	7.41417796124828	6.47932353965689
H	13.55668537154118	8.47589965211220	5.05991955803230
H	16.08707084308165	6.73045527247093	5.47085842579715
H	16.01011955785787	8.49091494138348	5.66800880189615
H	15.77202104406421	7.78893550881390	4.05708658752030
H	8.75765447079156	5.29718176427990	1.96569531010023
H	6.55596585723049	4.85943195306454	0.87181074488516
H	5.49917719204840	2.61163487401446	1.05658664091304
H	6.62605947361354	0.80177411134768	2.34565462163243
H	8.81091410690246	1.25045420237586	3.47048977143573

Table D.25. B3LYP/TZVP-optimized Cartesian coordinates of TOP monomer in the gas phase.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.52783817202448	6.34366329885436	4.65079861828102
O	11.47579904579221	7.10659447787938	5.24204518921004
N	10.37277182895416	5.00142658028190	3.71464265112128
N	10.18607815398487	3.96127721758786	2.93367855235833
N	11.38983961416150	3.47060125502486	2.49329142603840
C	11.67366544762652	5.21200242509883	3.82270748769050
C	12.37781684727754	4.24279371861243	3.05628444026244
C	13.71398554778350	3.93248639108204	2.76246375534839
C	13.97312286558764	2.87988007671950	1.91764043226574
C	12.91688495384117	2.13037004028690	1.35161761283541
C	11.61563321155031	2.42615042788723	1.64163549840426
C	12.18882377293523	6.33850890814662	4.66111901879002
C	14.17885354007833	7.40246681083225	5.43471032527450
C	15.67186728667477	7.22301095350197	5.30356053213205
C	8.90481533558486	3.44552209475739	2.53159315294904
C	8.10301816248321	4.22840526819617	1.70785291523212
C	6.85346726970519	3.74111015784142	1.34339228160936
C	6.43032185265943	2.49320190954869	1.79368390635230
C	7.24672159598162	1.72503488506255	2.62068054493532
C	8.49461046051957	2.20156489365099	3.00410532866553
H	14.49251864336274	4.53298926753640	3.20736399267067
H	14.99501613681060	2.62062217852729	1.67499211218417
H	13.12377140870883	1.30948704667094	0.67884810462409
H	10.75064845368712	1.91096004630178	1.25353961323389
H	13.83665520212279	7.31040758007230	6.46569146987452
H	13.83818282923105	8.36132817448333	5.04363575219725
H	15.99269679233722	6.25621350366442	5.69484055277125
H	16.17635987597626	8.00122969106145	5.87882575684544
H	15.99371941230963	7.31284107860060	4.26476109404071
H	8.44382561434682	5.19983726880252	1.37498563409825
H	6.21019078945232	4.33854865264572	0.71035622341827
H	5.45545188490815	2.12113612005159	1.50617876105645
H	6.90630264442582	0.76391182856681	2.98255998009914
H	9.12437139511447	1.62959978816151	3.67463929312983

Table D.26. B3LYP/TZVP-optimized Cartesian coordinates of TOP monomer in the gas phase in the excited state.

34
Coordinates from ORCA-job Triazolium-Ph

O	13.56098144490194	6.39682703436964	4.66303273431844
O	11.54491551491302	7.04384201590139	5.19433098287976
N	10.43381305598020	4.99807967008734	3.65070419235453
N	10.22721762924740	3.96511023899214	2.88893014331604
N	11.42939325820685	3.41908827176763	2.43670392323427
C	11.77119651120490	5.14668432643337	3.73284156307395
C	12.47491752078632	4.14344747828081	2.97745828376844
C	13.78431345964747	3.78451936854736	2.68958624811714
C	13.98272162554316	2.65280850554617	1.81559331933569
C	12.90985864880071	1.97281074327311	1.28456835275910
C	11.57588438262271	2.33091409484608	1.57481288577481
C	12.28522483662362	6.21689265949322	4.54667105785794
C	14.11582326278778	7.51744825952800	5.52417911056041
C	15.61180651999580	7.42060793844039	5.44322014725059
C	8.92919439024504	3.47158968171214	2.52179707770203
C	8.05842399639974	4.32928744556466	1.84456471177847
C	6.78378365698831	3.86379733501506	1.51972723851682
C	6.40339105797948	2.56122044778178	1.85749607062119
C	7.29024612920286	1.71713261516213	2.53371111479508
C	8.56485760895285	2.16889647216788	2.87855618726107
H	14.61689036380745	4.34432748320013	3.09707258262992
H	14.99721170255635	2.34361471794890	1.57054011138080
H	13.06606766804705	1.12781343695660	0.61432076730528
H	10.67663905730320	1.88569240750827	1.17080231758840
H	13.72204516994223	7.35291010865519	6.53609423402966
H	13.71661053711799	8.45185706197927	5.10826213169142
H	15.97852503365218	6.46672168347242	5.84278000436472
H	16.03790846921070	8.22810814803748	6.05655007024607
H	15.97094152633159	7.55599273188459	4.41526236510277
H	8.37545462105787	5.33925725354220	1.58417148919961
H	6.08822168090101	4.52048779366538	0.99558086108856
H	5.40707466690607	2.20282348525035	1.59582632438305
H	6.98609810277737	0.70642756302155	2.80737546609334
H	9.25799293735875	1.52814553796737	3.42559793962061

Table D.27. PBE/TZVP-optimized Cartesian coordinates TOP dimer in the gas phase.

56
Coordinates from ORCA-job Triazolium-Ph

C	1.14918201064942	0.76334807104665	0.66360894969216
C	1.17118823440760	-0.62611668264774	0.75743453705033
C	0.24306007455973	-1.37475179427741	0.02275110488558
C	-0.69881877468706	-0.75977543005791	-0.81108215221229
C	-0.72089596533101	0.62967241077822	-0.90478336022686
C	0.20717785731596	1.37829287803384	-0.17003371208984
N	0.17602976842339	2.80638290698997	-0.19196533665325
N	0.16991825291275	3.60893049583943	-1.32104493781794
C	0.15869918903445	4.92099513985997	-0.87295524570587
C	0.14828214354214	5.95364869088420	-1.82439304292519
C	0.17353870590472	5.62008286569040	-3.16738845242295
C	0.22376678308800	4.26461128623239	-3.57032158655693
C	0.22298647167276	3.24919112330702	-2.64236560844852
C	0.18978729080568	4.82047595258913	0.55279310888502
C	0.21325370891281	5.89081659547557	1.60862330292761
N	0.20414547761842	3.54173383622304	0.91272219444112
O	0.15327173670174	7.09991884465469	1.02583748764057
O	0.27686095673564	5.64176125249670	2.78986300376707
C	0.16633477401318	8.25773296352061	1.95120176534626
C	0.12372936014120	9.51537356040739	1.11968021882018
H	1.86167497546534	1.36836021240742	1.22446904773849
H	1.93903851002645	-1.11079404790496	1.36187598131829
H	-1.41134765869395	-1.36476802840144	-1.37190226588202
H	-1.48859160909594	1.11441878411807	-1.50938753993231
H	0.12938209004020	6.98232514318550	-1.46829251854215
H	0.16776402896355	6.40483505488050	-3.92444835488576
H	0.26879849215293	3.99924422566271	-4.62671022905740
H	0.27832400078575	2.18890284842359	-2.87769559930590
H	-0.70433202923716	8.15574761732678	2.61443996944222
H	1.07781878185815	8.18025266246494	2.56030731002277
H	0.13083142624286	10.38480631125852	1.79198109969371
H	1.00071656239934	9.59720031562367	0.46247660188984
H	-0.79186313093859	9.57105840990764	0.51471242627151
N	0.27349252502229	-2.80282465323854	0.04518204259085
N	0.27942801200049	-3.60525963439225	1.17444855779263
C	0.28680551152009	-4.91740924869968	0.72651672646666
C	0.29682235217323	-5.95002751712778	1.67796987917656
C	0.27493167146860	-5.61634909374662	3.02100548163553
C	0.22835538557335	-4.26074858372702	3.42388778398428
C	0.22968463021871	-3.24538315721261	2.49585751703748

C	0.25341042331216	-4.81704854099499	-0.69920153213684
C	0.22533102262718	-5.88742953802568	-1.75488630869546
N	0.24132091969082	-3.53836168707658	-1.05929152672418
O	0.28212039556866	-7.09666488149385	-1.17207920939136
O	0.16021972384768	-5.63826644821704	-2.93601866943921
C	0.26360925592779	-8.25451616597736	-2.09733033504229
C	0.30493485149749	-9.51220531220426	-1.26580428140049
H	0.31277970344984	-6.97875178017268	1.32186634217090
H	0.28049996040772	-6.40106072028251	3.77810962585457
H	0.18581030246721	-3.99517261344346	4.48032804548116
H	0.17697681900643	-2.18503326207757	2.73139869470965
H	1.13278815773516	-8.15493397646395	-2.76289519068742
H	-0.64926932952473	-8.17461294864011	-2.70401451298722
H	0.29361763817481	-10.38166711045907	-1.93801372394813
H	-0.57052546872907	-9.59155615153251	-0.60625021817043
H	1.22198204017366	-9.57035045079302	-0.66328835544482

Table D.28. PBE/TZVP-optimized Cartesian coordinates of TOP dimer in the gas phase in the excited state.

56
Coordinates from ORCA-job Triazolium-Ph

C	0.95299334475684	0.66505703916931	1.07218588929347
C	0.84942706844665	-0.71497824945872	1.23174620369948
C	-0.25772076892919	-1.40822015517624	0.70514403486181
C	-1.25840283531860	-0.69745284096147	0.00049122357132
C	-1.16237300921220	0.67410541800039	-0.15646132294544
C	-0.05613074386040	1.34519946821495	0.38621229763235
N	0.04245193145548	2.77522433540855	0.20857348438926
N	0.61721518248582	3.36234118347077	-0.92810376594785
C	0.51656420336498	4.73502513843873	-0.80881740472209
C	1.02896220223535	5.53582162789604	-1.84213365002821
C	1.60967776225534	4.90422842108459	-2.93092366018794
C	1.68543682236110	3.49436236603294	-3.00121246922400
C	1.18170996833931	2.71255829369148	-1.98553357477217
C	-0.14172878317763	4.88959544676448	0.44547635070269
C	-0.54697512719710	6.11467045108364	1.13656423168722
N	-0.40145961969176	3.70523651255061	1.00348111812647
O	-0.32163355439573	7.25580913627126	0.60290865569538
O	-1.13040159322409	6.07938898188508	2.27173669572959
C	-0.74894255461109	8.57521862587086	1.31552727809907

C	-0.33661465777031	9.70204621627166	0.42348530710887
H	1.81398866372285	1.19517411144586	1.48193782777595
H	1.64853381189684	-1.24671524802986	1.74494226421315
H	-2.09775846801469	-1.24990880677845	-0.41843427864663
H	-1.94899758742012	1.21577166435940	-0.68459497304326
H	0.96412180275726	6.62046195576723	-1.77819350143477
H	2.01458203537377	5.50233725027221	-3.74858438052940
H	2.14207238375689	3.00391244922658	-3.86101883407030
H	1.18620589132795	1.62396059488937	-1.94269748930946
H	-1.83424587833462	8.48127325579067	1.44866626866373
H	-0.22845649160385	8.55290898989442	2.28089872026831
H	-0.63965612481051	10.63478783672193	0.92493854417761
H	0.75078126829379	9.74436379504078	0.28456263389273
H	-0.84900494422991	9.67261874495110	-0.54556967995050
N	-0.45632674741923	-2.78382786451515	0.87109374771843
N	0.58688566443230	-3.61198912436331	1.39895835815871
C	0.53769571455346	-4.81665554343166	0.66132536290906
C	1.32364841053385	-5.89161078865374	1.11912785210073
C	1.95633313219099	-5.79597622930873	2.34055522841806
C	1.74143778169497	-4.65065326387255	3.17500279575647
C	1.00938348019008	-3.59354044093629	2.72048531773807
C	-0.46767195636408	-4.62588406608230	-0.31346276375412
C	-0.92495282618904	-5.55052165079367	-1.39695026021933
N	-1.06192117954032	-3.43575268370390	-0.21349486378125
O	-0.27915324443644	-6.73586433390392	-1.31791770623373
O	-1.76187661933685	-5.25324262722285	-2.22266631673301
C	-0.65825747720814	-7.73186773645249	-2.32860964107454
C	0.15056427347491	-8.98118070931395	-2.06714765251532
H	1.35384085091824	-6.79842224328702	0.51881890140180
H	2.55436320422997	-6.62618443126325	2.71280158064980
H	2.13316630092743	-4.62289595475323	4.18983239558965
H	0.75882350928666	-2.71623021414811	3.31113561374197
H	-0.45904081452031	-7.30238184756957	-3.32091896707779
H	-1.74021262522884	-7.90667101556268	-2.24402323498942
H	-0.11589442040101	-9.74596536268731	-2.81018106462806
H	-0.05931596838772	-9.39295490771852	-1.07013879984452
H	1.22909895557072	-8.78762997051596	-2.15332092810811

Table D.29. BLYP/TZVP-optimized Cartesian coordinates of TOP dimer in gas phase.

56
Coordinates from ORCA-job Triazolium-Ph

C	1.18463708609288	0.76455069450381	0.61822944356347
C	1.20767456641667	-0.62884165719311	0.71366760230914
C	0.24168776226737	-1.37895890451032	0.02488414260842
C	-0.74065700905637	-0.76102066166266	-0.76458618558973
C	-0.76371801686076	0.63236734838162	-0.85998885176503
C	0.20225994698615	1.38248004066765	-0.17118498277902
N	0.16840052239474	2.82149722283707	-0.19617751613541
N	0.17109092925228	3.63598259663623	-1.33549612775883
C	0.16203807562017	4.95232037161244	-0.87961966020489
C	0.16074766191101	5.98904114470154	-1.83220268348580
C	0.19024805784077	5.65995530805693	-3.17964234269878
C	0.23223254213242	4.30236609919166	-3.59012629636884
C	0.22366934674577	3.28072436740538	-2.66470716260260
C	0.18219793286177	4.84998016392610	0.55172580818620
C	0.19905349957069	5.92449399921747	1.61534240167622
N	0.19068604428428	3.56853488703384	0.91418082721671
O	0.14815705728773	7.14352121045186	1.03431486649971
O	0.25497454550923	5.67097646217625	2.79834796427837
C	0.15912143836340	8.32034246109163	1.96787287125646
C	0.13000119161870	9.57983433585276	1.12381963839107
H	1.92413916880745	1.36486915899175	1.14273600281938
H	1.99726140252469	-1.11373649491485	1.28416235649888
H	-1.48012026107907	-1.36135550135193	-1.28912585346850
H	-1.55323156641820	1.11729163328021	-1.43056363723009
H	0.14450835879917	7.01491159878566	-1.47727613002498
H	0.19170170750478	6.44631325015529	-3.93151874932019
H	0.27611435571134	4.04308031861102	-4.64564193980385
H	0.26805275195653	2.22407760153035	-2.90502880080698
H	-0.71829560721852	8.22142129562301	2.61721039060070
H	1.06447538496398	8.23691055362408	2.57958353493398
H	0.13528639290190	10.45073906425646	1.79222034335835
H	1.01193930606512	9.65238030149719	0.47493891001529
H	-0.77845253939272	9.63536464878882	0.51100621655619
N	0.27561197898289	-2.81796604021227	0.04984081463373
N	0.27267794574197	-3.63252984636775	1.18913599589061
C	0.28076314832449	-4.94884593757382	0.73316755297421
C	0.28190951309392	-5.98564575477111	1.68565750694879
C	0.25324623043142	-5.65665763725831	3.03314288471533
C	0.21227038538808	-4.29907936291494	3.44374377947240
C	0.22102158672006	-3.27736374084701	2.51840036579060

C	0.26010653454521	-4.84640757387790	-0.69816412545146
C	0.24233516329592	-5.92085156695417	-1.76183757714773
N	0.25222748631745	-3.56494440710109	-1.06053875674896
O	0.29286787207260	-7.13993953415006	-1.18093348271700
O	0.18669872849401	-5.66719407504312	-2.94483091379743
C	0.28137806976025	-8.31667353263731	-2.11460548581018
C	0.31065220665492	-9.57625119633793	-1.27068881664529
H	0.29741035608143	-7.01149141317522	1.33062849550016
H	0.25170490888215	-6.44307741727076	3.78495434894107
H	0.16906249983438	-4.03983494029929	4.49929820378854
H	0.17740895967181	-2.22072768785508	2.75888682198552
H	1.15854881888779	-8.21781274552399	-2.76428756122002
H	-0.62420850900495	-8.23304283328205	-2.72594187791580
H	0.30481995153655	-10.44709090021546	-1.93917021932703
H	-0.57096659395928	-9.64865739851311	-0.62135627650181
H	1.21940972188159	-9.63204837707349	-0.65835307808326

Table D.30. BLYP/TZVP-optimized Cartesian coordinates of TOP dimer in gas phase in the excited state.

56
Coordinates from ORCA-job Triazolium-Ph

C	0.89920859881640	0.64591129834235	1.09541314783025
C	0.78093569702791	-0.73596441783869	1.27437954851278
C	-0.34736640771063	-1.42246061025027	0.78255345676512
C	-1.36617508025545	-0.70313454879859	0.10869657592252
C	-1.25833760565693	0.66942618343864	-0.06841646995487
C	-0.12155131224488	1.33204525932730	0.42835871857370
N	-0.00260702418172	2.76911092663177	0.22137488967385
N	0.50900197271024	3.34177542477911	-0.97328752759283
C	0.44921517190516	4.72234946833533	-0.85161011584427
C	0.90249591285784	5.50873124826239	-1.92761950055113
C	1.38698625620722	4.86108663235119	-3.05759970040696
C	1.42604035669669	3.44613244110748	-3.12838896359605
C	0.98093210330499	2.67592760620893	-2.07312026657083
C	-0.11189291077132	4.90317363494309	0.45130473749499
C	-0.43187214872273	6.14715160751289	1.16688732324468
N	-0.36222410449289	3.72658984095352	1.03794084181608
O	-0.21830690089710	7.28443349243556	0.61451355236533
O	-0.92933102304405	6.12671807892378	2.34941914818351
C	-0.55568880636178	8.65978374428316	1.35867539152006

C	-0.17595854293403	9.75929877200755	0.41071006379038
H	1.77644893721107	1.16798343117863	1.47295385423685
H	1.58202553915982	-1.27178905674647	1.77401428846278
H	-2.22764759388363	-1.24214785590043	-0.27384857415658
H	-2.05394194693856	1.21503791211432	-0.57329776269628
H	0.86769091861027	6.59212081582728	-1.86458304310551
H	1.74163891003976	5.44696995698466	-3.90353280349059
H	1.80529179126817	2.94565190410886	-4.01641111508420
H	0.96587020939758	1.59086833455182	-2.03312037820582
H	-1.62620575716791	8.58749759830489	1.56988704115148
H	0.04233384357190	8.61686756234675	2.27290967279915
H	-0.41093422582494	10.70729391439845	0.91860195398081
H	0.89643417906516	9.76318766836202	0.18873847208777
H	-0.76176460205977	9.72986814153729	-0.51374590061982
N	-0.55734929579497	-2.81155399267809	0.96021969816246
N	0.52437627839627	-3.62909875521717	1.49038313987118
C	0.53471549374307	-4.81828909239346	0.70958004835700
C	1.32613141009995	-5.89149401383931	1.17090197876569
C	1.89930595403330	-5.82525290467741	2.42731478654783
C	1.61608430598515	-4.71237185739327	3.29281218389222
C	0.88000474531947	-3.65457987688371	2.84123978077314
C	-0.43062611818945	-4.62217156121220	-0.31069390436361
C	-0.79126435162779	-5.51747914197599	-1.46544076315949
N	-1.07756022024757	-3.45872623335899	-0.20296363612144
O	-0.17190773844513	-6.72362661370852	-1.35836963512830
O	-1.53732226598508	-5.18288772698446	-2.36465039671345
C	-0.46085304685587	-7.70037084537802	-2.44037556415083
C	0.31884303842901	-8.96831926135452	-2.13756582124245
H	1.40619724755504	-6.77651267282669	0.54844697445850
H	2.49653045670515	-6.65322074017602	2.79884241200469
H	1.95969447798494	-4.71451517078556	4.32312629338917
H	0.58516016541124	-2.80386257302747	3.44548136652953
H	-0.16566368900030	-7.24337662931083	-3.39208326229646
H	-1.54396097745581	-7.86782810639832	-2.45771009601579
H	0.11749764343502	-9.70864308879225	-2.92328900695697
H	0.01609749187572	-9.40457004580918	-1.17748819922157
H	1.39996358992680	-8.78046450584314	-2.11896393391745

Table D.31. PBE0/TZVP-optimized Cartesian coordinates of TOP dimer in gas phase.

56
Coordinates from ORCA-job Triazolium-Ph

C	1.16822690720797	0.75080619025413	0.62936409845677
C	1.18660398057532	-0.63028891164158	0.71545847634443
C	0.23643927449929	-1.36383849537048	0.01568958468320
C	-0.72429996761216	-0.74726413544383	-0.77559290298328
C	-0.74270028156714	0.63382180010817	-0.86164523808983
C	0.20747995763943	1.36737076111532	-0.16188744418922
N	0.18307921034560	2.78837993566394	-0.18156097924215
N	0.17785241342965	3.57877658672486	-1.28994008061350
C	0.16563492719888	4.87713658596213	-0.85587789637686
C	0.15817427871626	5.90374140301002	-1.80285135425410
C	0.18384375040136	5.56577616655638	-3.13257840463626
C	0.23102676852549	4.21469080530414	-3.52655791235166
C	0.22911287728185	3.21420573314886	-2.60014649670745
C	0.19159413773986	4.77796503474487	0.56194042820581
C	0.20916125556704	5.84874263321265	1.60703771266420
N	0.20516889980667	3.51370927011231	0.91198873058850
O	0.15356568044621	7.03959732198213	1.03254622738673
O	0.26943436234881	5.60481628015788	2.77448504665185
C	0.16439708642988	8.18839800554986	1.93377009055261
C	0.12238152041451	9.43375944626364	1.09631238630444
H	1.89705716950717	1.34647370861708	1.16503191014735
H	1.96349834536657	-1.12011574131912	1.29095864765002
H	-1.45309689875064	-1.34294216446674	-1.31128893271899
H	-1.51952162547896	1.12368795049853	-1.43721462918669
H	0.13961834215783	6.92684304622197	-1.45308039322700
H	0.18020317665988	6.34095504045037	-3.88930393487989
H	0.27397000918850	3.94658101525457	-4.57478642260065
H	0.28053981246571	2.15941632646734	-2.82790911683971
H	-0.70070281487891	8.09841895717444	2.59343473271690
H	1.06821198641176	8.12342940278188	2.54216082843782
H	0.12785333629680	10.30424848397232	1.75486500951307
H	0.99417573036603	9.50646170514418	0.44307240033880
H	-0.78604333443171	9.47980718391448	0.49265043446478
N	0.26083543160050	-2.78483678341491	0.03532952051292
N	0.26608135142325	-3.57534444513949	1.14364637017123
C	0.27734825471596	-4.87367157540277	0.70945038631228
C	0.28492886160742	-5.90039348090334	1.65628085260678
C	0.26028346465115	-5.56258939935126	2.98607270645897
C	0.21402309867694	-4.21153014301713	3.38022450044232
C	0.21588002074381	-3.21093059314731	2.45393087627508

C	0.25062032086426	-4.77436389187829	-0.70834252380809
C	0.23199016628520	-5.84502084523250	-1.75355436170318
N	0.23756359212406	-3.51007732756696	-1.05826612950406
O	0.28723171479556	-7.03597077024234	-1.17923930902649
O	0.17140471031104	-5.60090595603364	-2.92094685762988
C	0.27558744470568	-8.18464981305690	-2.08061734126757
C	0.31808310752870	-9.43013243945618	-1.24336267998099
H	0.30276247412100	-6.92345498370595	1.30635403513609
H	0.26404272634942	-6.33786295349025	3.74270043369255
H	0.17187788116690	-3.94350040384508	4.42850661172360
H	0.16511438863250	-2.15616530329683	2.68193360678398
H	1.14019140152290	-8.09468207475563	-2.74093538647674
H	-0.62868664251759	-8.11948362841007	-2.68830218599882
H	0.31194753255163	-10.30053011538555	-1.90203079019497
H	-0.55319776840347	-9.50279354518333	-0.58943257421796
H	1.22698519226837	-9.47639586021098	-0.64043936651709

Table D.32. PBE0/TZVP/RIJCOSX-optimized Cartesian coordinates of TOP dimer in gas phase in the excited state.

56
Coordinates from ORCA-job Triazolium-Ph2

C	1.07960798231100	0.77743973441827	0.70462814869463
C	1.12565871715315	-0.59705118590753	0.80511718931560
C	0.28668768678515	-1.38043988264152	0.00213790465450
C	-0.56566737000424	-0.77296635572762	-0.93022420243856
C	-0.61489415540619	0.60134968495290	-1.01896776110976
C	0.20830075268029	1.37255724358770	-0.20189423539409
N	0.12822297955569	2.78908831916322	-0.21218931263422
N	0.26832678331578	3.59118314467883	-1.30358912584134
C	0.19144011057395	4.88722672394572	-0.86667389546448
C	0.29514179149446	5.91942940914190	-1.80431413679079
C	0.48728215903238	5.59005660317445	-3.12097187835338
C	0.59239958814194	4.24030109490480	-3.51339167173210
C	0.48220669105276	3.23504980772236	-2.60030130610567
C	0.03282958604265	4.77665096638958	0.53772190227564
C	-0.09497404783571	5.83536021354052	1.58269214663752
N	0.00168220212986	3.50622771190471	0.87732757654338
O	-0.01171962133717	7.03619495620817	1.02737912082727
O	-0.24515444389465	5.58447647936652	2.74066910613759
C	-0.10056906101255	8.16847606160514	1.94158217142924

C	-0.05005472357224	9.42647974925641	1.12301265013175
H	1.72367911334963	1.39168282090482	1.32162438703049
H	1.86006455985279	-1.04665497037078	1.46061527875583
H	-1.20017207518247	-1.38854560428263	-1.55353682592816
H	-1.32931797540455	1.06460467976140	-1.68996090826254
H	0.22640659047683	6.94054620366761	-1.45573423620740
H	0.56978488980202	6.36938160975091	-3.86846732295054
H	0.76688378472146	3.97979847427877	-4.54947444548560
H	0.56554322233597	2.18093919579048	-2.82082956722884
H	-1.03135357094693	8.07260291046322	2.50296274914796
H	0.73184481966604	8.09244190174377	2.64383999385398
H	-0.11709973927050	10.28802087005947	1.78983886368482
H	0.88792929525201	9.50562823778420	0.57012733649901
H	-0.88626452255197	9.48024990643137	0.42332576720817
N	0.31563544013488	-2.77119997756428	0.03529999377257
N	0.46928071467061	-3.56332096068641	1.16817467972575
C	0.42025014238331	-4.87846883335761	0.76335913044689
C	0.56717656905864	-5.87499000914337	1.74294747643498
C	0.68771933045565	-5.49920355496354	3.05382449271222
C	0.65374020824999	-4.14067616164077	3.42400744718634
C	0.53901110658629	-3.17655346693862	2.46416706257199
C	0.17695184639833	-4.80816805574962	-0.60856746246373
C	0.01823008933091	-5.87936824179232	-1.53565709090285
N	0.09942504926928	-3.52460857145025	-1.05165457202176
O	0.11763345994398	-7.10425105961523	-1.15990419186058
O	-0.22084485414802	-5.66812076203327	-2.75921309507379
C	-0.04487889290217	-8.22094555553136	-2.14638711332830
C	-0.00575689673175	-9.49480396563155	-1.36575910933530
H	0.56453856076190	-6.91369815749253	1.44758606954959
H	0.79318184622679	-6.25690625114998	3.82089628032872
H	0.71133088364732	-3.84147620414751	4.46156430070983
H	0.48062559037681	-2.11945419729074	2.67277172194784
H	0.77855335568346	-8.12456403355850	-2.85475662084707
H	-0.99790981350961	-8.05965035373531	-2.65166344729258
H	-0.11508943987549	-10.32572941935226	-2.06668703985896
H	-0.82625017854500	-9.55312715791029	-0.64988387931802
H	0.94760288322723	-9.62022076493179	-0.85104149398370

Table D.33. B3LYP/TZVP-optimized Cartesian coordinates of TOP dimer in gas phase.

56
Coordinates from ORCA-job Triazolium-Ph

C	1.20383724197091	0.75127738784624	0.58451891733568
C	1.22264607134949	-0.63386564015858	0.67114202911168
C	0.23678787907517	-1.36829729132397	0.01598624262197
C	-0.76134334188501	-0.74778623712353	-0.73128659245906
C	-0.78011903833341	0.63735740404897	-0.81795345261700
C	0.20572120777507	1.37179481581144	-0.16277453739960
N	0.17836207038367	2.80217469773981	-0.18373897479340
N	0.18192184584596	3.60300378605118	-1.30216000002403
C	0.17110764611721	4.90661764883132	-0.86029353320918
C	0.17146158934921	5.93655273403542	-1.80857730608123
C	0.20153505076101	5.60248798269074	-3.14333326893859
C	0.24208291542408	4.24955265809235	-3.54420526209728
C	0.23311204607245	3.24236552084636	-2.61965020195682
C	0.18685243671587	4.80572126583558	0.56273090177859
C	0.19849537997186	5.87947003321846	1.61375962214598
N	0.19508087416171	3.53763527657135	0.91560257366337
O	0.15013579446156	7.08019548785953	1.03922964412746
O	0.24828125354455	5.63253788395873	2.78503420754237
C	0.15496195400605	8.24538543838652	1.94738347353701
C	0.12138233343847	9.49237793717812	1.09994102976320
H	1.95817257073431	1.34298792572766	1.08538925649543
H	2.02077956176158	-1.12575254426683	1.21221223490848
H	-1.51572870176206	-1.33950112593935	-1.23208086788915
H	-1.57828839007542	1.12925510679464	-1.35895937998291
H	0.15478350239603	6.95811580456655	-1.45909467612101
H	0.20421610244473	6.37976848441501	-3.89618180354794
H	0.28452737192740	3.98710821019527	-4.59275951686666
H	0.27548802062311	2.18973651280355	-2.85201006466497
H	-0.71629833613651	8.15491735794239	2.59626315339060
H	1.05392748680773	8.17757221958244	2.55993864741370
H	0.12241054087334	10.36496435974852	1.75550953782405
H	0.99828204248881	9.56079104151335	0.45381366782809
H	-0.78146331700843	9.53653480505437	0.48852568072311
N	0.26414565858311	-2.79867701053128	0.03703483329004
N	0.26050330985281	-3.59942872198166	1.15550815956356
C	0.27133839461698	-4.90307523618356	0.71372568665108
C	0.27083262045464	-5.93294601504546	1.66208074047552
C	0.24070022929353	-5.59879646187805	2.99680905316855
C	0.20017485336867	-4.24583556333961	3.39759217015808
C	0.20922899485664	-3.23870614351595	2.47297014592875

C	0.25572050596869	-4.80227291967225	-0.70930278085239
C	0.24410593372402	-5.87608890122751	-1.76026625507030
N	0.24758805213585	-3.53420475743163	-1.06225703766794
O	0.29264083290986	-7.07676649221420	-1.18563360640630
O	0.19481082604917	-5.62924009794945	-2.93157544290395
C	0.28831641030542	-8.24203696012687	-2.09368466009951
C	0.32188154839091	-9.48893912499085	-1.24610852335434
H	0.28749122643253	-6.95453082145940	1.31266864020316
H	0.23792297874026	-6.37603250598497	3.74970120788147
H	0.15764576680964	-3.98332200730173	4.44612603555674
H	0.16681648975194	-2.18605976550169	2.70526227592369
H	1.15976450974180	-8.15145161618324	-2.74229176929253
H	-0.61046696042768	-8.17445249988489	-2.70652891877986
H	0.32116285983761	-10.36158994263529	-1.90158830749079
H	-0.55517441066389	-9.55742955526730	-0.60020157117618
H	1.22457670398697	-9.53289282822682	-0.63445645726845

Table D.34. B3LYP/TZVP/RIJCOSX-optimized Cartesian coordinates of TOP dimer in gas phase in the excited state.

56
Coordinates from ORCA-job Triazolium-Ph2

C	1.01528456776465	0.77759497854176	0.77447800399343
C	1.06347279001224	-0.60003901674693	0.86336057429861
C	0.22563887384583	-1.37523213982113	0.05752318341791
C	-0.66024215230712	-0.77606883045906	-0.83896965531572
C	-0.70682795665155	0.60189406624038	-0.93902135937108
C	0.14597270852346	1.38972486607572	-0.14641396070933
N	0.11408198533109	2.78696550749745	-0.18343967123562
N	-0.05341621901622	3.59165276245449	-1.33017364494695
C	0.01092366075695	4.91503532096866	-0.91513012482868
C	-0.14580273069137	5.91674485582238	-1.89809998689345
C	-0.28342205439226	5.54553020169753	-3.20906176715620
C	-0.25999779631794	4.18042373978544	-3.58831392563734
C	-0.13899676374098	3.21005029013299	-2.63418049177257
C	0.27173031084204	4.83977632669705	0.45191145483509
C	0.44939114889521	5.91243839718071	1.39234926823386
N	0.34828255571080	3.55514755171389	0.90010016271112
O	0.36171410704427	7.14130447113159	1.02373162191527
O	0.69657025732660	5.69534698206301	2.61776632934136
C	0.54446751191397	8.28660559963144	2.01663489965037

C	0.47820180195313	9.55715672543594	1.22253893540175
H	1.66202146991044	1.38867622003626	1.38697273831234
H	1.78526340846231	-1.06344660758228	1.52416967848653
H	-1.31321081546383	-1.38724011011866	-1.44732454838279
H	-1.44644641739620	1.05168349555775	-1.58590837909886
H	-0.13697001085808	6.95427235623820	-1.60293370275023
H	-0.39576196411601	6.30500722228144	-3.97167591537258
H	-0.33129369991369	3.88865377955300	-4.62591469916840
H	-0.09254416329166	2.15550651692064	-2.84946250004575
H	-0.26282888313340	8.17987892883359	2.73909418152379
H	1.50995404648710	8.12297677629861	2.49196830493535
H	0.60534880813267	10.39183875702480	1.91598339685703
H	1.27819123184088	9.61278910966497	0.48443119449309
H	-0.48899902817164	9.67613710111006	0.73434995221544
N	0.30898993744063	-2.79972498398975	0.06996831023813
N	0.15165312062840	-3.61323586342366	1.16868932944271
C	0.23608605871042	-4.91407883145010	0.72462545835581
C	0.12182662714253	-5.95046358919979	1.66140509752824
C	-0.08834071364741	-5.62624096772474	2.98049277129259
C	-0.19884279290755	-4.27507652035461	3.37949678608399
C	-0.07802016851646	-3.26269237211406	2.47035109101291
C	0.41402707698005	-4.80142681692890	-0.68192212152086
C	0.55350341504692	-5.86252071172947	-1.73080220729538
N	0.45067844890482	-3.52697813986202	-1.02310400905482
O	0.48760568545207	-7.07376320607674	-1.17436377835433
O	0.69766848701084	-5.60946880885530	-2.89329780682158
C	0.59046020324995	-8.22023073594140	-2.09470751185973
C	0.54282579975608	-9.48178392936060	-1.26910654562349
H	0.19718327189171	-6.96940005459394	1.31248325841561
H	-0.17908674729427	-6.40829819167887	3.72247829526531
H	-0.38493473775095	-4.02084641073076	4.41396551686647
H	-0.16102062000661	-2.21133090106827	2.69611745064474
H	1.52435770889900	-8.11675494973635	-2.64672878018579
H	-0.23824496625638	-8.14843134889889	-2.79955263119983
H	0.62043283083508	-10.34370895850136	-1.93409453358076
H	-0.39788847609993	-9.56604073700034	-0.72247594668548
H	1.37416896123933	-9.52800817264178	-0.56375204090121

Table D.35. PBE/TZVP-optimized Cartesian coordinates of TOQ monomer in gas phase.

40
Coordinates from ORCA-job S0gasOpt

O	13.63973271549595	5.46680891666839	5.32088673936176
O	11.71396736824273	5.35610609352554	6.55043157831788
N	10.46256088644604	4.18683406913550	4.30906417604563
N	10.19081541908225	3.63979891994140	3.14409382281117
N	11.32330615517196	3.59996848229973	2.33333399070202
C	11.75472883895036	4.51969974982253	4.31616717220280
C	12.34132595225627	4.17001579140725	3.07092155572647
C	13.63519191762449	4.29294628173485	2.51557838815671
C	13.84764761479957	3.83310338512938	1.24516216744022
C	12.80305637527116	3.21754595916851	0.48941177662060
C	11.48949449086247	3.07105375481984	1.04010794469422
C	12.34218203431257	5.15860574646871	5.53566424510055
C	14.32756850292592	6.11001770940417	6.45495370937959
C	15.74632152356298	6.40384164010396	6.02896458994993
C	8.84499546888706	3.26115785453419	2.78533503456240
C	8.05486492498550	4.17053059771297	2.07710115854210
C	6.73960072196813	3.81273608676599	1.77634411710369
C	6.23554065159365	2.57511569862538	2.19142132707234
C	7.03789821091915	1.68900358751583	2.91766522525785
C	8.35711887589356	2.02872442567072	3.22486830582611
C	13.03870757276737	2.71625631894488	-0.81260503814637
C	10.47595318325200	2.42405921072044	0.31284811187079
C	10.75737610445425	1.94425182676866	-0.95897403045357
C	12.03448392262859	2.09238453003147	-1.52932492822569
H	14.41073343223076	4.75243818539676	3.12375713866980
H	14.27671358815447	5.42219034125846	7.31067545205354
H	13.76792820415352	7.01962634523054	6.71462466477621
H	16.28960737381323	5.48412531265762	5.77124664555713
H	16.27760667800746	6.88407601278472	6.86319298646255
H	15.77477925644562	7.09106274686458	5.17188397143428
H	8.45726169596006	5.13841090795635	1.77580302400715
H	6.10309778950101	4.50998022792757	1.22992827908192
H	6.63562274684757	0.73345061232225	3.25538058224561
H	8.99503567813999	1.35677524341668	3.80075451356289
H	9.48242781032069	2.28958670939390	0.72813781412357
H	14.03771738098427	2.82868624216313	-1.23769514288638
H	9.96757249566698	1.44150101113511	-1.51847466505569
H	12.23191198825808	1.71076249161808	-2.53127148437889
H	14.83444333644240	3.92119970897763	0.78782023656264
H	5.20485273571984	2.30500484697633	1.95785718786450

Table D.36. PBE/TZVP-optimized Cartesian coordinates of TOQ monomer in gas phase in the excited state.

40
Coordinates from ORCA-job TDDFT_PBE

O	13.63816768226675	5.33551345491013	5.40990239015256
O	11.91241856047701	4.78329081478426	6.61914911622167
N	10.70091929562541	3.56700384751854	4.42599933876662
N	10.36275579304836	3.08826835236862	3.21033691228303
N	11.35604311842947	3.49256582849147	2.25576107567729
C	11.89249305012621	4.15071631314897	4.29358283969648
C	12.38466202689012	4.13856316274152	2.96768241053119
C	13.52342689619891	4.55767173338736	2.27392116074653
C	13.64486772689825	4.23234254923601	0.93566849992856
C	12.69193099560150	3.38452756796562	0.27718184568955
C	11.53404133012523	2.92895894235275	0.98745133274975
C	12.48249398571231	4.76281801503247	5.47275323319415
C	14.24773942153028	6.00559298521924	6.63263315564890
C	15.53318878577947	6.63701363073680	6.18276898003306
C	8.96038818529767	3.08650863209055	2.85308089122158
C	8.46800056928962	3.97077357766546	1.89019402544631
C	7.10580206558126	3.93536039217378	1.58280231961869
C	6.25505272852609	3.04266897575932	2.23966327578735
C	6.76540218697424	2.17413313482470	3.21005806193928
C	8.12681857825948	2.18220437720096	3.51557228711384
C	12.88420896181857	2.91345486054060	-1.03614747698146
C	10.65452306080941	2.00516894698108	0.41659063315990
C	10.89413965234480	1.55187773409741	-0.88668162723954
C	11.99358991744337	2.01102292628780	-1.61568506423960
H	14.30676135715120	5.10514356525239	2.79168469070143
H	14.38817991213895	5.20975062905957	7.37545448808944
H	13.50552897343962	6.73230486799255	6.98848965426621
H	16.24224559829015	5.88931070618249	5.80523169271029
H	15.99306382840878	7.12355415123623	7.05579043686948
H	15.36192782962061	7.40824083557817	5.42098452831351
H	9.13450382171114	4.67265063074258	1.38740730187053
H	6.70895920281189	4.61681187668747	0.82859918337000
H	6.10413581401257	1.47534918887900	3.72420556183090
H	8.54366705051407	1.50191922134573	4.25940460064135
H	9.79405638174716	1.63695838271233	0.97023169492931
H	13.76242126684323	3.25410233589091	-1.58740313559796
H	10.20691498176025	0.82927422028271	-1.32836469899264
H	12.16259953178838	1.65585548578137	-2.63232580644698

H	14.52139361652238	4.55785764087148	0.37519993869596
H	5.19231788118575	3.02233908898959	1.99421259560355

Table D.37. BLYP/TZVP-optimized Cartesian coordinates of TOQ monomer in gas phase.

40
Coordinates from ORCA-job S0gasOpt

O	13.64400700030767	5.47133539818545	5.33155918755573
O	11.71917613445836	5.30631290090417	6.57263086635174
N	10.46669794166448	4.14745764065424	4.31932939365108
N	10.18491565394188	3.59670546811289	3.14733892794919
N	11.32988641790880	3.57821574512717	2.32116037783489
C	11.75854709963842	4.49252824412982	4.32161819209988
C	12.34582635756615	4.15617124201178	3.06667407874273
C	13.64137166908004	4.29364952272149	2.50872074060112
C	13.86193731707488	3.84088366247527	1.23404125542048
C	12.82400452657336	3.21560515973467	0.46902425711179
C	11.50630940886553	3.05242269093590	1.01889848974709
C	12.34614747638512	5.13214445265484	5.54998220641756
C	14.34252559920287	6.12558385809181	6.47638804538243
C	15.76068887156877	6.44041123863394	6.03432333665620
C	8.82161375602415	3.23968129853823	2.78692476885009
C	8.03519564602607	4.17407949271029	2.10040789296401
C	6.70585479506518	3.84427447538160	1.81380440251946
C	6.18254548938096	2.60838469609426	2.22174543118866
C	6.98170030207841	1.69545887989402	2.92433479277233
C	8.31507685298337	2.00731316184761	3.21640406698080
C	13.06703591152923	2.72037911418070	-0.83892157302492
C	10.49902777633136	2.39700578444303	0.28346594966852
C	10.78798057258864	1.92338673543081	-0.99240007858961
C	12.06822675259740	2.08766540704559	-1.56096818762566
H	14.41151789957509	4.75623393204294	3.11617838737143
H	14.30698480727634	5.43149814573805	7.32379432072971
H	13.77328221710710	7.02460896923585	6.73909012212330
H	16.31124204425353	5.52975864987353	5.76705610644713
H	16.29247284782819	6.92259851465689	6.86556128247268
H	15.77110463788506	7.13081173338276	5.18148156547399
H	8.44811160330803	5.13753674611768	1.80792669935629
H	6.07616874943521	4.55925195388672	1.28732583427331
H	6.56793302171594	0.74478236774257	3.25468720220549
H	8.94507109421704	1.31641693068286	3.77348535946598
H	9.50909423856175	2.25167060365098	0.69492391902029

H	14.06264294882516	2.84467437968026	-1.26232530540321
H	10.00640488909720	1.41603851143443	-1.55483000734489
H	12.27057935123378	1.71262575518273	-2.56195729566288
H	14.84741064728620	3.94223189394186	0.78249415300731
H	5.14543129655311	2.36164822581030	2.00166317923850

Table D.38. BLYP/TZVP-optimized Cartesian coordinates of TOQ monomer in gas phase in the excited state.

40
Coordinates from ORCA-job TDDFT_PBE

O	13.62417244345430	5.36800620447266	5.41090818689148
O	11.89236113051471	4.80671033855206	6.62200944411522
N	10.69815382445668	3.55173603673552	4.43839929224835
N	10.36076012900494	3.02575978915041	3.22336591735968
N	11.36353385528080	3.45126630852156	2.25167841668600
C	11.88635412096785	4.14889171114210	4.29410946953950
C	12.38471196513594	4.11793869824805	2.96765243928832
C	13.52718881617191	4.54183537161209	2.27176617596838
C	13.66408971484525	4.20469029294272	0.93643703576219
C	12.71533330801698	3.34943504289872	0.26968037542817
C	11.54910805833653	2.89188882896221	0.97585049317865
C	12.47082321209832	4.78244214260522	5.47316152620746
C	14.24835469011518	6.06423503141049	6.64524996616082
C	15.54814496008374	6.67244909093785	6.18188134388080
C	8.94169580694459	3.06302296639133	2.86372105981525
C	8.43971698389837	4.05714417575519	2.01580054169223
C	7.07311617898827	4.05381205294936	1.70922572414485
C	6.22679831268297	3.08151152115113	2.25841215519172
C	6.74584656467887	2.10144049373892	3.11599401626296
C	8.11314887847615	2.07959830314109	3.41547676774471
C	12.91462366328101	2.88015770155271	-1.04740030070703
C	10.66748446136542	1.97250954637265	0.39130843608388
C	10.91346733745896	1.52235332371046	-0.91460461084857
C	12.02346374101768	1.97990564067189	-1.63650422805875
H	14.29980937849936	5.10200134451955	2.78710198566749
H	14.37244014094864	5.27637113102809	7.39449681424280
H	13.50953301182450	6.80097740967643	6.97567199122840
H	16.24682743146399	5.90907482755247	5.82196808624139
H	16.00976550776180	7.17025722204905	7.04616868357800
H	15.38742012145952	7.42801875142507	5.40478234037375

H	9.09960589337275	4.81406466850879	1.59632239972444
H	6.67089210625050	4.81713458335394	1.04483232802227
H	6.09066953745525	1.34459843772785	3.54367547870223
H	8.53451296380453	1.31684698560083	4.06746219782807
H	9.80271407268496	1.60792724300776	0.93468203171722
H	13.79315229035042	3.22057862069485	-1.59365958338908
H	10.22696003823165	0.80691659424816	-1.36357629007007
H	12.19770620286804	1.62765469827550	-2.65089333976682
H	14.54198818177782	4.53198381798864	0.38383844437443
H	5.16530258697066	3.08629663371660	2.01658913148922

Table D.39. PBE0/TZVP-optimized Cartesian coordinates of TOQ monomer in gas phase.

40
Coordinates from ORCA-job S0gasOpt

O	13.61221208553297	5.47144652324517	5.30454065719385
O	11.69455333986789	5.42943417619912	6.49042259489257
N	10.45967233905636	4.23439321421205	4.27523620124243
N	10.19795627856164	3.68349919656732	3.12775113666879
N	11.31688119636055	3.62170337216039	2.34367924204839
C	11.73937135595834	4.55158790194058	4.29494719006677
C	12.32803727881039	4.17910985293853	3.06899469998631
C	13.62514574837864	4.27544723107395	2.53026246435260
C	13.83679997676428	3.79921735672270	1.27988607795061
C	12.78852530833145	3.20128123411549	0.52511111044003
C	11.48052512215079	3.08989353189498	1.05948518673978
C	12.32523403963468	5.20021526209741	5.50150979278028
C	14.29779581746939	6.11137995318418	6.41394876574235
C	15.72360735698346	6.34799506226151	6.00198581890383
C	8.86495934326118	3.28126742281761	2.78027067416326
C	8.04965508547905	4.18123589996978	2.11092058830227
C	6.75093180993059	3.79848734341962	1.81307867227034
C	6.29001591277492	2.54402724798225	2.19406500176198
C	7.11785675826627	1.66621951664604	2.88268162118167
C	8.42037155469349	2.03300699336395	3.18625969563675
C	13.02719955266292	2.69378808392430	-0.76371194470426
C	10.46326042712227	2.47559523526435	0.32485971181940
C	10.74413682371219	1.99011750243813	-0.93329563733451
C	12.02348570412202	2.09899180089813	-1.48555876261966
H	14.39621401939973	4.72747384263869	3.13641373790474
H	14.21934997778681	5.45473836399953	7.28230587692913
H	13.77283639736571	7.04031641630748	6.64428707099883

H	16.23400860514186	5.40996234502102	5.77540807428851
H	16.25783348117458	6.82948036264475	6.82357508156680
H	15.78207175632309	7.00556570606912	5.13249862002148
H	8.42178748708278	5.16050662754600	1.83417944805948
H	6.09430514981297	4.48668182167993	1.29437195651580
H	6.74749894723220	0.69702560328634	3.19359641315040
H	9.07790284770170	1.36728688589858	3.73317929182237
H	9.46704210984952	2.37130201461667	0.72233936749603
H	14.02671330014962	2.78017281597364	-1.17460633002560
H	9.95185150160760	1.51407066679387	-1.49846478465922
H	12.22050056960139	1.71173494699364	-2.47752860723339
H	14.82155181672875	3.86071559411665	0.83107260894058
H	5.27209344015594	2.25306865407655	1.96308395873819

Table D.40. PBE0/TZVP/RIJCOSX-optimized Cartesian coordinates of TOQ monomer in gas phase in the excited state.

40
Coordinates from ORCA-job Triazolium-Ph3

O	13.39075304332041	5.13073496712637	5.26826352951066
O	11.81065229725210	4.22612372556727	6.59444311719161
N	10.77865706991937	2.93753167424881	4.36306054255013
N	10.42851042095002	2.57818390687227	3.01641723881021
N	11.35307558674686	3.22991074493135	2.15577206805636
C	11.78578818854565	3.77196165899519	4.26772212796795
C	12.21376402489861	3.97422764908203	2.93628592053926
C	13.27176092158106	4.63674777634349	2.30662073075311
C	13.49187596403309	4.40329250262894	0.97768817151104
C	12.72413064694266	3.44662727274532	0.24183020085378
C	11.65069453621090	2.77909751634475	0.88817599461369
C	12.31588728068861	4.38780384761484	5.52167000474628
C	13.98486195101377	5.79661377984066	6.40984051984461
C	15.13660754230561	6.62391133776787	5.91015248720413
C	9.09912737503656	2.97933724480744	2.86260043676721
C	8.63335871814709	3.73081646332458	1.77488492903190
C	7.29997555161302	4.04561377208991	1.71773142609620
C	6.40048950113486	3.62025093463829	2.73655616472074
C	6.86116598732759	2.86927177839145	3.81727785728622
C	8.18785598669728	2.53948493600673	3.89172950630865
C	13.01717008702286	3.08713133511854	-1.07816765371259
C	10.92837478085471	1.77857624020576	0.24592257965494

C	11.26047254228655	1.44594923265372	-1.05751385430938
C	12.29231883217486	2.10341161928553	-1.72312042218751
H	13.91334126748999	5.27435462213400	2.89713800910955
H	14.30110882329517	5.03240492032069	7.12260918071843
H	13.21466762452638	6.40501101065496	6.88761472713253
H	15.90171125915847	6.00010530896589	5.44459764330952
H	15.59520101005154	7.14786187692124	6.75084094318445
H	14.80209341994802	7.37322606986411	5.19015061086634
H	9.31386036186676	4.05935564304483	1.00386394170629
H	6.91601554923555	4.62496655365151	0.88672906011182
H	6.17536050211308	2.54129385729635	4.58774107114845
H	8.58913629578762	1.95482389234160	4.70728181728555
H	10.14741002941711	1.24288219609544	0.76930673322532
H	13.83825590022550	3.58319381234075	-1.58315999030837
H	10.71382663898368	0.65542579318210	-1.55763392253438
H	12.53784876284817	1.82949460787021	-2.74149711929230
H	14.31231444057302	4.89931517909796	0.47252432717795
H	5.35227090077585	3.88311632258726	2.65909168734961

Table D.41. B3LYP/TZVP-optimized Cartesian coordinates of TOQ monomer in gas phase.

40
Coordinates from ORCA-job S0gasOpt

O	13.61894481207748	5.47058576126525	5.31668930239975
O	11.69702554526885	5.38777451338653	6.51339813623860
N	10.46140212060821	4.20255970752469	4.28491961924126
N	10.19105504462637	3.65064174682447	3.12906498296659
N	11.32212917153912	3.60374508513171	2.33263027968648
C	11.74271953805430	4.52910256112875	4.30128986554658
C	12.33275588566570	4.16839227753666	3.06632494462021
C	13.63036501942717	4.27982841902878	2.52363547819980
C	13.84872060258598	3.81363223786970	1.26656198407264
C	12.80683360386512	3.20608941672267	0.50413107498803
C	11.49503882605289	3.07506202377431	1.04017416276891
C	12.32815241881545	5.17636086462979	5.51617534166208
C	14.31510410586214	6.11732315464938	6.43677051385668
C	15.73933301076888	6.37692803725786	6.00860911189376
C	8.84346284973832	3.26441961439153	2.78034051284819
C	8.03513777569923	4.17874849438024	2.11410745699793
C	6.72323472960410	3.81868262300031	1.82904728189599

C	6.24064655893223	2.57135949092357	2.21956741645315
C	7.06171536970617	1.67796241471945	2.90322558295105
C	8.37735848977069	2.02222282749503	3.19349257590685
C	13.05107247212995	2.70497472204759	-0.79101259186363
C	10.48467650237220	2.44666768983246	0.30117984487799
C	10.77167485873827	1.96756654296729	-0.96188801387744
C	12.05246771923159	2.09770731660397	-1.51580152512580
H	14.39795155793345	4.73429566334388	3.12985807024304
H	14.25269939858458	5.45004243780079	7.29645580540666
H	13.77862031330836	7.03563110334197	6.67618574780496
H	16.25910830751312	5.44800778137914	5.76820724153913
H	16.27558201092547	6.85835934965616	6.82882384358560
H	15.77929777745078	7.04223050714429	5.14446271799445
H	8.41978660213314	5.15054620963640	1.83262701857802
H	6.07527044911869	4.51746291359922	1.31574999857872
H	6.67744532170867	0.71732706814252	3.22026199542851
H	9.02537882156614	1.34601523659352	3.73647674723079
H	9.49239794388644	2.32654492131698	0.69885900296336
H	14.04754479309979	2.80576660681847	-1.20300903855900
H	9.98718116413687	1.48191424787598	-1.52775709616234
H	12.25385865821163	1.71718405587528	-2.50838351281141
H	14.83265428647823	3.89031242985639	0.82110841793140
H	5.21594718580422	2.29946550752706	2.00048204504247

Table D.42. B3LYP/TZVP/RIJCOSX-optimized Cartesian coordinates of TOQ monomer in gas phase in the excited state.

40
Coordinates from ORCA-job Triazolium-Ph3

O	13.39120996595315	5.14535042468311	5.28574201877326
O	11.73207413231615	4.32148962480537	6.58362458845271
N	10.74863590755311	2.97329299776288	4.35862191301857
N	10.40887791195963	2.58137486431586	2.99724411277906
N	11.33902160767421	3.25825619784664	2.13200312467103
C	11.75632495096115	3.81347662832609	4.25988428772942
C	12.19739242982211	4.00567782106141	2.92818207435369
C	13.26352036991315	4.67052420287525	2.30780370833568
C	13.50812378188105	4.43176791921169	0.98076257288212
C	12.75796594028309	3.46224627045849	0.23592198032936
C	11.67048958905077	2.79290533935984	0.86788148247422
C	12.27478292347402	4.44249010710251	5.51931984730291

C	13.98693739076930	5.81488827333794	6.44414929965053
C	15.17281978310131	6.61325233021414	5.95751430958611
C	9.06400282644308	2.97279171873909	2.84371870470388
C	8.58749237128593	3.75884977996580	1.78918869793999
C	7.24127037762148	4.04689592808804	1.73361007988278
C	6.33833741508830	3.56170633157098	2.73120312323938
C	6.80996229711155	2.77478680033689	3.78194549886300
C	8.14810486999444	2.47155017857481	3.85053437191698
C	13.08317697900244	3.09024244344478	-1.07773876159867
C	10.96986788025592	1.77872789446181	0.21605717957773
C	11.33400837498689	1.43311691429570	-1.07946170596059
C	12.37909796940016	2.09199117973911	-1.73014694590548
H	13.89554949611113	5.31042407822846	2.90340152840271
H	14.27049083008154	5.04684266037297	7.16438627372695
H	13.22185188447293	6.44364535352519	6.89970139320942
H	15.92893583114194	5.96850596800577	5.50714641852612
H	15.62955177172376	7.12747134040187	6.80517067747854
H	14.87021954001711	7.36817964306287	5.22997111159280
H	9.26047496353696	4.13331687252146	1.03514849846260
H	6.85116002954713	4.64777362139155	0.92249637513080
H	6.12934776995132	2.40117893936054	4.53444118300047
H	8.55372122805324	1.86394002032930	4.64494400243465
H	10.18085687190734	1.24331803019110	0.72421195012197
H	13.91101132783773	3.58593964973095	-1.56947961386680
H	10.80431720690881	0.63440443608311	-1.58245680471037
H	12.64935531764189	1.81011808285450	-2.73892234418467
H	14.33417031838976	4.92917647107206	0.48894341346851
H	5.28723918977497	3.80755624529013	2.65637271820866

Table D.43. PBE/TZVP-optimized Cartesian coordinates of TOQ dimer in gas phase.

68
Coordinates from ORCA-job GeoOpt

O	13.26194818483197	5.82259366022609	5.14348180174077
O	11.17710460332819	6.75004173350757	5.34105377617046
N	10.28429968374865	4.89232710790395	3.43080758867500
N	10.19326006860702	3.99639305672528	2.45131290098948
N	11.44795169342524	3.47185523007421	2.14730837979914
C	11.56066355491825	4.99000072695665	3.77960716427360
C	12.34235426138753	4.08420174597835	3.00248514477666
C	13.70939525110666	3.74159357711695	2.95746551238500
C	14.12221750271121	2.81608702517375	2.03330371634235

C	13.21718739500781	2.26988317486135	1.07385019353691
C	11.83574206339506	2.64702177687953	1.07902108005413
C	11.95845051945424	5.96880559254557	4.84994731819283
C	13.78369628152595	6.72458344521768	6.19655874771049
C	15.24989002771619	6.41991568154229	6.38179135094205
C	8.91340005271258	3.39419075407964	2.19367044376566
C	7.83726217793370	4.26497866087629	1.98653001245411
C	6.54719633865382	3.74822639213927	1.89103787130559
C	6.36185401221660	2.36373327657834	1.96929196096912
C	7.43969347480594	1.49206342272598	2.16169924313599
C	8.72557608257709	2.01155549746942	2.29284403032499
C	13.66417592697823	1.40222605135371	0.04714983050328
O	1.40331854971337	0.26112631760260	4.11132061477807
O	3.37573554631899	-0.51119317655413	4.98234691628052
N	4.71530304596303	1.04387748031805	3.06597599634774
N	5.04710942493841	1.78432574647859	2.01226280273339
N	3.90825060350244	2.21806181389809	1.33714221853962
C	3.39141252610812	0.96069667189910	3.10761260447613
C	2.82873587021420	1.71435483702205	2.03519413009645
C	1.51564810767045	2.00291105769305	1.61016734321344
C	1.34579301046285	2.76548049198866	0.48306802448248
C	2.46029935020596	3.18953334235136	-0.30212585003750
C	3.79675680030229	2.86127503045814	0.09388766930532
C	2.74006490281866	0.14216853768558	4.18789806475020
C	0.63725859514608	-0.49612131553855	5.12870410664841
C	-0.83114320360270	-0.26842698244024	4.86677358595624
C	10.97479259805481	2.24127500800081	0.04663846795691
C	4.88474208936891	3.13350692245029	-0.75084392430567
C	11.46486384403632	1.41087775464695	-0.95259297242463
C	12.80253818711710	0.97081791782037	-0.94410855999571
C	2.28116931753508	3.87635629077952	-1.52785286248277
C	4.65676484077858	3.78656960279927	-1.95476702726987
C	3.35984208169209	4.18008095992579	-2.33703697397797
H	14.38277387865965	4.22714379515266	3.66081496181494
H	13.19294104063453	6.54316489031586	7.10545791876518
H	13.60128564811066	7.75655910971597	5.86632207168015
H	15.41017559801390	5.38181355026978	6.70261140682180
H	15.65147290649574	7.07753261748387	7.16574666115729
H	15.82285926848603	6.60915069426510	5.46358939743079
H	8.00399636863345	5.34195397555862	1.95667436243381
H	5.70180554453202	4.42761365731600	1.78023418940827
H	7.26414730196245	0.42114196648945	2.26556583528434
H	9.55893377704631	1.33868560631582	2.49583336828178
H	0.68902421713141	1.60798350039562	2.19671059743744

H	0.95523766045132	-0.13206551812457	6.11551459721373
H	0.92646013715272	-1.55257764889106	5.03992553609021
H	-1.09819049062092	0.79380700328255	4.95133681680965
H	-1.41444107111261	-0.81952619874417	5.61748897462227
H	-1.12801271106129	-0.64050982230679	3.87669418001299
H	0.34195349025338	3.02479867399667	0.14292297353597
H	9.94457193557621	2.58554525655708	0.00861493725113
H	14.71277153221918	1.09968837170365	0.04210349436906
H	5.89193256068795	2.82155781548803	-0.48717746683468
H	10.80166000774400	1.10587488201532	-1.76320667853982
H	13.16204049871382	0.30987361895556	-1.73369703670461
H	1.26700114635783	4.14082772434219	-1.83149388034081
H	5.49888364362686	3.98633532557289	-2.61877306177392
H	3.20631784479671	4.70063250769397	-3.28267744337886
H	15.16993322212166	2.51340996796274	1.99108572403339

Table D.44. PBE/TZVP-optimized Cartesian coordinates of TOQ dimer in gas phase in the excited state.

68
Coordinates from ORCA-job GeoOpt_S1dimer

O	13.36308987789630	5.81202570080064	5.30740781375844
O	11.32011588934994	5.96325293262004	6.09187065617960
N	10.35442585366300	4.36689326427416	4.02829755228990
N	10.17880716623109	3.63003440955524	2.97227518591507
N	11.40679215439703	3.44243903230895	2.29776548222639
C	11.64748584726443	4.68684576980011	4.09667133834524
C	12.37154471480137	4.11599532012951	3.01903392898960
C	13.71892847241029	4.10725603978171	2.60976261765131
C	14.04129631079519	3.40929699613917	1.47339656054068
C	13.05601612823155	2.71878691954687	0.71435433870346
C	11.67622907130428	2.72697110330022	1.11633392068530
C	12.12168859128929	5.51479029696748	5.20324866423170
C	13.87997969882165	6.68838360455506	6.48213557326604
C	15.36397318933886	6.78843513931792	6.32691803419241
C	8.89235006284157	3.06500125186443	2.61796562095992
C	7.96985253853092	3.85430854831245	1.93105108148924
C	6.71077677240419	3.32766021572648	1.64368560113166
C	6.38887205599939	2.01614362530821	2.03784935334196
C	7.33611242197685	1.24257662216797	2.74703565016357
C	8.58439424629664	1.76678199966601	3.04450281878990

C	13.41022942789644	2.01431538039982	-0.46176930574008
O	1.37058280381827	0.37046337967577	3.98659934394658
O	3.18195509507960	-0.87355707141337	4.60536339926687
N	4.70194421427763	0.53310910763364	2.73189635291308
N	5.15307335252539	1.40905997020939	1.73294741938067
N	4.05951918614355	2.22984770788337	1.30141000079961
C	3.40871230952534	0.79059923221986	2.90271773467581
C	2.94164598459897	1.82013296895142	2.05113511327094
C	1.69404344032999	2.39969719897367	1.75151139649722
C	1.57600736871258	3.19119003598009	0.63435247455846
C	2.65992665141462	3.31463542945552	-0.30717516714761
C	3.92140615024200	2.72085272141348	-0.00776603937006
C	2.66606391558881	-0.00966978559824	3.92913987970933
C	0.54201288443413	-0.35223870950608	4.96087401772174
C	-0.86246161044081	0.19754887081987	4.86679377856129
C	10.70564993124772	2.06150233867839	0.34966764874371
C	4.94332498492940	2.67752639103875	-0.96106272339923
C	11.10464075355840	1.38631544832613	-0.79606331693260
C	12.45119708375378	1.35623731226253	-1.20658596817586
C	2.50287624516260	3.91980791252618	-1.56962401319478
C	4.73959701004002	3.27015701171717	-2.20962676396873
C	3.53077119937425	3.90815351760714	-2.50767518299831
H	14.46508631759394	4.64194842451472	3.19159832885719
H	13.56047719446476	6.16285497394129	7.39114426871199
H	13.35204253216444	7.64447786495952	6.37365262653576
H	15.85473681454044	5.81084528122785	6.40773297602098
H	15.73351445310788	7.41097654683131	7.15683438732549
H	15.64957179058172	7.28626255882113	5.39175661895400
H	8.21171169664232	4.87622710286093	1.63716252384978
H	5.97614538939610	3.94912609440805	1.13548215782869
H	7.05324026959862	0.24835001879169	3.08758165358568
H	9.30347363446887	1.17217114581732	3.61064651096827
H	0.84707387259632	2.16245489334205	2.39083041667810
H	0.98514717788680	-0.20659416495055	5.95681469224121
H	0.59112180176164	-1.42340271539947	4.71985961130655
H	-0.89137530851921	1.26931736339578	5.10732654460637
H	-1.50483783171275	-0.32748402084189	5.58718662028122
H	-1.28650833006746	0.04532472688158	3.86428433496320
H	0.62034991575131	3.65391183289430	0.38797436519449
H	9.66010263974553	2.06373850765803	0.63815914267996
H	14.45781659872875	2.00639975267318	-0.76808365523942
H	5.87104774951979	2.14906771145478	-0.74441847197342
H	10.34911980548260	0.86845140173223	-1.38800944777811
H	12.73316802755364	0.81632827688988	-2.11094948816994

H	1.54221799428560	4.37737063142836	-1.81171600642482
H	5.52891866522262	3.22102323828583	-2.96094044705138
H	3.38419392086129	4.37306082435354	-3.48255462474535
H	15.07710396428795	3.37847378863144	1.13012736882410

Table D.45. BLYP/TZVP-optimized Cartesian coordinates of TOQ dimer in gas phase.

68
Coordinates from ORCA-job GeoOpt

O	13.28457729422742	5.71993515931745	5.26950567836707
O	11.17421852729477	6.58842299241072	5.54323428147990
N	10.29397349812616	4.84172593903219	3.51801864275367
N	10.19828254130611	4.01131286446902	2.46826318856383
N	11.47588352964768	3.52079977631490	2.11963491102708
C	11.57521444165243	4.93860022545813	3.85820064274019
C	12.36769252472807	4.09984468599313	3.00939485335078
C	13.74612591731398	3.79792611261172	2.92287088203862
C	14.17546560650302	2.95053414802461	1.92943444138795
C	13.27281070150985	2.43566196106835	0.94475701127359
C	11.87461040824482	2.76667419520406	0.99382255178164
C	11.96814351759648	5.85648515396028	4.99406751406596
C	13.81396515425356	6.56988889848911	6.38708606717081
C	15.29489918312915	6.27392538854478	6.53071824837946
C	8.91327243500219	3.39423275577175	2.20047360567303
C	7.82891839150604	4.25881304167165	1.98986856280983
C	6.53851958624352	3.73146881484842	1.89463409375718
C	6.36129795348100	2.34306296314276	1.97351493855455
C	7.44837047103493	1.47670340233319	2.16132141746217
C	8.73404556856534	2.00708943985919	2.29565372306065
C	13.73094468535882	1.64352675661222	-0.14265763414661
O	1.34611725186433	0.38536030789098	4.21134877567923
O	3.32019832235826	-0.33793017292295	5.13992844214936
N	4.68289705097373	1.08316790681698	3.12937983101062
N	5.03868930150908	1.75007054375800	2.02085115297022
N	3.88882285712095	2.15952285913033	1.31035309165852
C	3.35594231888426	1.01359275478257	3.16674988403075
C	2.80185769555354	1.70222376315175	2.03925311027429
C	1.48864893464340	1.96152669321631	1.58438233317082
C	1.32207206253201	2.64273794314962	0.40230365881712
C	2.44116910034862	3.02110789450786	-0.40652849036643
C	3.78179247992802	2.72763611282922	0.02184545083690
C	2.68979099137412	0.26630742199154	4.30011600106248

C	0.55416171636393	-0.31082491102027	5.27944823794698
C	-0.91699066989794	-0.09336810475692	4.98044698942857
C	11.00692487758259	2.38223268206178	-0.04606863239277
C	4.87781895967901	2.96538483755232	-0.82887455978821
C	11.50745737835737	1.62649658354610	-1.10228685750367
C	12.86436995107196	1.23809045027806	-1.14449912397685
C	2.26790031364066	3.62966171194254	-1.67886727274728
C	4.65511848519593	3.54172634938212	-2.07578889390531
C	3.35410357127316	3.89446388995415	-2.49646010901712
H	14.41564799669909	4.25035111792815	3.64680655263055
H	13.23932379351725	6.31530896998314	7.28488556177889
H	13.61080208879667	7.61362031210468	6.12265314059917
H	15.47347741336063	5.22086158733451	6.77953987907252
H	15.69768836711673	6.88497063615112	7.34909486528161
H	15.84926704662800	6.53045927061162	5.61924966119703
H	7.98588797839080	5.33441031208995	1.95670813306851
H	5.69101625551626	4.40393723617380	1.78459657932979
H	7.28349988667210	0.40620193733173	2.25915845503423
H	9.57016111719656	1.34134738507912	2.49664652109698
H	0.65916370769268	1.60848259355631	2.18787264502976
H	0.86193974347945	0.11935228146563	6.23909730838838
H	0.84240779452591	-1.36756300027563	5.25815979480956
H	-1.17929341261825	0.97158814350618	4.99461937701333
H	-1.51135724782090	-0.59478154523217	5.75530692004821
H	-1.20020267196749	-0.52386563889609	4.01201945327884
H	0.32052559665946	2.87096101147427	0.04159307272047
H	9.96637763466194	2.68536538868312	-0.04600939993089
H	14.78623920064204	1.37811792686201	-0.18361156280954
H	5.88326392356932	2.68606975957780	-0.53610764204273
H	10.84023138454418	1.33996974617587	-1.91315221113114
H	13.23106531378594	0.63907587027279	-1.97576427047794
H	1.25731917475887	3.86595709944888	-2.00783970976634
H	5.50039333696398	3.71361917535524	-2.73970918904556
H	3.20563557487938	4.35285599455386	-3.47185655894549
H	15.22953431726705	2.68775145830440	1.85529686288226

Table D.46. BLYP/TZVP-optimized Cartesian coordinates of TOQ dimer in gas phase in the excited state.

68
Coordinates from ORCA-job GeoOpt_S1dimer

O	13.34571696751142	5.59535708290177	5.15399974542185
O	11.39778205684912	6.79268454009084	5.29391623689669
N	10.38771045689623	5.29123359122119	3.16074518711304
N	10.16964333939604	4.33265005046387	2.12973436048645
N	11.32969063220593	3.45504162742396	2.06905522167950
C	11.59301246648693	5.03737736058270	3.66903309604068
C	12.23648468602624	3.93777604336052	3.04336611696093
C	13.51268407412733	3.33736424430647	3.11467592804841
C	13.90190028543083	2.46264931524311	2.12514553515359
C	13.08845194941379	2.25747187863620	0.94673955920676
C	11.79311601820361	2.86068878661040	0.87127769979814
C	12.07482377408479	5.91698364237663	4.79276836936674
C	13.93299272727744	6.40664884683419	6.25134987015265
C	15.35294647256948	5.91549782496552	6.48018959983935
C	8.87562894618799	3.74624608328760	2.20146650444723
C	7.79365758867042	4.59791192975775	2.53652799853548
C	6.49732787098915	4.09934914424277	2.59184070968381
C	6.27811797555943	2.74448554612044	2.29621620254822
C	7.32817766789381	1.88065590467872	1.97174275844102
C	8.63219128825361	2.38626977588669	1.92444989847949
C	13.55656814607240	1.55828014544225	-0.18856607201414
O	1.22126983648029	0.40922159302025	4.33135047157080
O	3.01306434525152	0.52452676557844	5.60499723290861
N	4.47476807367407	1.69052298447362	3.53796884493986
N	4.92406058316854	2.20895031286967	2.42611571940862
N	3.88236305533994	2.21242152529775	1.43949800319113
C	3.19687077369395	1.34224376061992	3.35649417524811
C	2.76400769862294	1.65493246286254	2.03667739532906
C	1.55485819979614	1.52041087945713	1.32054882046449
C	1.51920646681525	1.95958018077964	0.01786077169602
C	2.66212511638935	2.53331027046835	-0.61374921202114
C	3.90737058792509	2.67314947591665	0.10196414857491
C	2.45450923051793	0.73839887909738	4.46854781202490
C	0.39254045446325	-0.24300347533094	5.52106640106310
C	-0.99730836053051	-0.46724551558278	4.99939968336487
C	11.04147852177952	2.81092709376057	-0.31103545568567
C	5.03533482560684	3.23126909063982	-0.52936896713818
C	11.55108598060940	2.12638153163608	-1.42153044091516

C	12.79707025050165	1.48578679256380	-1.35637798745468
C	2.60608005727198	2.97967584807093	-1.96129979523193
C	4.92715994337231	3.65342014221745	-1.85081876456671
C	3.71900726346952	3.53177749058280	-2.57144930274317
H	14.17506434075204	3.62617642381870	3.92346746073402
H	13.29755605084015	6.28262135666844	7.13628393733780
H	13.89758661914766	7.45823473583785	5.94508635451268
H	15.36894639715350	4.86078420452667	6.78118469305501
H	15.80937245138175	6.50436420531391	7.28659726691164
H	15.96909678978714	6.03916231140444	5.58074996707081
H	7.99633019053112	5.63455370444207	2.78745555848545
H	5.67547553960061	4.75503877219983	2.87375683629861
H	7.14915501715070	0.82474986246711	1.77700663995245
H	9.45326099265408	1.71334153795834	1.69989253255886
H	0.68863863441065	1.08142047165204	1.80290182235610
H	0.46107011761692	0.49001475234361	6.33010960756394
H	0.94190871208630	-1.15845814560909	5.75890234631751
H	-1.49205861715422	0.47105931212724	4.72655196352414
H	-1.57509264341892	-0.91941864907813	5.81953709373013
H	-1.01385296140177	-1.16904439387229	4.15837976112327
H	0.59867062658246	1.87009616738605	-0.55719452933469
H	10.09196507815328	3.33436012649433	-0.37695425218057
H	14.54165994281655	1.09652127869469	-0.14662133340621
H	5.97545744667271	3.34022768059646	-0.00670250452618
H	10.97660321118989	2.10371302382527	-2.34568778399609
H	13.18211648946457	0.95324744900884	-2.22328068848440
H	1.66814379240414	2.87641732659619	-2.50455419982304
H	5.80003127051519	4.08767900964536	-2.33418447476911
H	3.66609705600750	3.87219813110932	-3.60328941936010
H	14.88135936273096	1.99190511301009	2.16617614403421

Table D.47. PBE0/TZVP-optimized Cartesian coordinates of TOQ dimer in gas phase.

68
Coordinates from ORCA-job GeoOpt

O	13.22070038348844	5.77517383390602	5.15187655498735
O	11.14719142892732	6.62311301781869	5.43505585364610
N	10.25865399353028	4.83556627193230	3.48178576367404
N	10.16879054716910	3.97718068975457	2.49130573629810
N	11.40504347254194	3.49436503468546	2.15656155034799
C	11.52235343324281	4.94903813840519	3.81134315519636
C	12.29911369202523	4.08608751742585	2.99853452178723

C	13.66452238148008	3.76715913544110	2.92396051357482
C	14.06487453789787	2.87472072876288	1.98077346033988
C	13.14929884731911	2.33975439252437	1.03322947006465
C	11.77851489424124	2.69706008109790	1.07419557568813
C	11.92586076258420	5.89116735975456	4.90168312928429
C	13.74830765579370	6.63727327278206	6.20354201429373
C	15.21929132137010	6.36292354624940	6.33244991300418
C	8.89753212111473	3.37893387078142	2.22628664864144
C	7.82953841033429	4.24201633014419	2.02195669595084
C	6.54911739309042	3.72547336760225	1.92372078055612
C	6.37075687620827	2.35103964288228	1.99511472840532
C	7.44093361221361	1.48651370001622	2.18253448254218
C	8.71646395780677	2.00585080681466	2.31863792801025
C	13.58577303021264	1.50495304482269	-0.01229492944286
O	1.43662216126138	0.31403697176134	4.12010196454320
O	3.37146499889174	-0.41073121990298	5.03101295071948
N	4.72612166392854	1.06760370101278	3.08956383072512
N	5.06295480848173	1.77381599159892	2.03368765422756
N	3.94916665041789	2.18702341076949	1.35364934338553
C	3.41845986706752	0.98571791374297	3.12666493551905
C	2.87109704148324	1.71084910309178	2.03833800084657
C	1.56808528512397	1.99296211858766	1.59824963671779
C	1.41845010902028	2.72961299074148	0.46620917836982
C	2.54127352015064	3.13327149341563	-0.30745008394429
C	3.85660556175252	2.80839247035620	0.10809314839312
C	2.75247897567195	0.20577321497768	4.21650362353481
C	0.66170279144540	-0.40087062674249	5.12860306630883
C	-0.79477005046624	-0.18865520674589	4.82998811887996
C	10.90658123942142	2.30565242321235	0.05672280693058
C	4.95122554877186	3.05836847913196	-0.72105992817974
C	11.38225653617694	1.50667009259081	-0.95983062349561
C	12.71690093741650	1.08595577051964	-0.98785532411328
C	2.37967905225987	3.79502325626491	-1.53884240401890
C	4.74383833054869	3.68803944594231	-1.92853014036481
C	3.46175005229515	4.07792745344383	-2.33299669952402
H	14.33941361937836	4.24077525568023	3.62241443062986
H	13.20028240376689	6.41491115691178	7.12104370564965
H	13.53867737780440	7.67086639351310	5.92278093878761
H	15.40913382002976	5.32409742369084	6.60756031277076
H	15.63025356397124	6.99698312101296	7.12024484501280
H	15.75231528196563	6.59336951526971	5.40809667756010
H	7.99297806888146	5.31236358515885	1.99511595276826
H	5.70671491410156	4.39733726474248	1.81354928398794
H	7.26871019860315	0.42144023695096	2.27727777863340

H	9.54814613042204	1.34061277551935	2.51619765074931
H	0.73881451952274	1.61591219514637	2.17912633842585
H	0.94878801142880	-0.00912723291303	6.10599641528755
H	0.94830126649038	-1.45292389884342	5.08451938810395
H	-1.06120120076941	0.86901456674088	4.87036651604230
H	-1.39042704548879	-0.71101749296973	5.58067700105579
H	-1.06500361285614	-0.59064010570204	3.85190118162761
H	0.42742154689042	2.98634807953807	0.10983515425419
H	9.87895400204295	2.63576767356304	0.04410979419345
H	14.63064850681412	1.21782121939067	-0.04362025014615
H	5.94642073479346	2.74774670152367	-0.44129838606880
H	10.71094309231479	1.21089430545633	-1.75732798884021
H	13.06625533684238	0.45045480941458	-1.79264009233105
H	1.37743774053881	4.05703837070631	-1.85777599714927
H	5.58964940967822	3.87260248048532	-2.58021991295483
H	3.32429293460085	4.57929986318400	-3.28311944984007
H	15.10761174651886	2.58649592745428	1.91021698947910

Table D.48. PBE0/TZVP/RIJCOSX-optimized Cartesian coordinates of TOQ dimer in gas phase in the excited state.

68
Coordinates from ORCA-job GeoOpt_S1dimer

O	12.95515830653315	6.02086948693143	4.67202647666225
O	11.30151841590983	7.33750934855409	3.89252269797971
N	10.44640875616089	5.22432198286521	2.32956450323612
N	10.30716095185654	3.98679126407703	1.68574787347391
N	11.39998945040672	3.16062537742459	2.02970233371753
C	11.51477137096295	5.12717733602324	3.07592641974409
C	12.16251040638017	3.87184664660101	2.94192584587684
C	13.34208117586765	3.29652035481948	3.40891021602266
C	13.77989443895320	2.13506366027348	2.82572909439339
C	13.13461581136873	1.59014169510873	1.67570091072284
C	11.94098404132108	2.19428692538277	1.19513125004674
C	11.89949464254135	6.30207027395920	3.92069450021175
C	13.42751796230876	7.09463412572105	5.53120845499803
C	14.52495267844970	6.54282647002595	6.39584495943166
C	9.03134919383439	3.47356283261248	1.75315855873403
C	7.94744104552627	4.39369166887566	1.72712953168363
C	6.65880295994627	3.93838930946334	1.73660941404504
C	6.41884609275545	2.55504313994370	1.77656908552872

C	7.48188722133469	1.63753891929751	1.84407948352131
C	8.77566834841949	2.08984347595407	1.82689107755715
C	13.67117267917266	0.52644327477546	0.93831669940024
O	1.77783503417719	-0.02125230037704	3.95916054077669
O	3.79228237956392	-1.00316557578984	4.26045981739469
N	4.91121334431279	1.03625450299875	2.73967428046724
N	5.12662222777969	2.02840442638966	1.90073470833886
N	3.94313271754050	2.54629217823480	1.43975217349138
C	3.61515658495112	0.91015517255666	2.89631989775600
C	2.94712575872075	1.88208086018388	2.10480433681442
C	1.60977336962355	2.21522922258886	1.89985405238163
C	1.31864186959287	3.19164755338132	0.97958408338050
C	2.32952686447711	3.77115533150665	0.18401688652870
C	3.69610291628827	3.40074552845047	0.35580498723719
C	3.07961579078313	-0.16539231779478	3.79690026059749
C	1.12601437011761	-1.00473204962385	4.82775455785708
C	-0.34674522694422	-0.71823424335322	4.82521614880024
C	11.35687783301531	1.78926813637509	0.00217412552869
C	4.65722984352374	3.82282286781446	-0.55391215366819
C	11.93309587711786	0.74512802632897	-0.70660214173151
C	13.07504629195353	0.10326750587838	-0.23241643126527
C	2.00186743750007	4.66139263958162	-0.86264787046483
C	4.28322335863282	4.66626111503034	-1.58592235707404
C	2.96413912243938	5.11054183550850	-1.72901395499294
H	13.89841358652416	3.80069320486394	4.18595220014110
H	12.58056544951351	7.45834424137048	6.11509775895732
H	13.77093820721636	7.90831080635514	4.88919261880948
H	14.16158026614672	5.72844499278633	7.02515848065163
H	14.89550244444834	7.33349488475748	7.05085388746478
H	15.36371243752666	6.18455470524004	5.79668737023807
H	8.16675005863916	5.45274133365058	1.73827375058697
H	5.84390489414713	4.65088475233856	1.76272811143998
H	7.27223670275806	0.57797979425435	1.92037795862371
H	9.59051146799422	1.38280180123457	1.88269112512505
H	0.85410951558393	1.69138492434531	2.46882255418507
H	1.57102929368860	-0.90679744060869	5.81941075989620
H	1.36314397325532	-1.99592288343450	4.43876406298248
H	-0.56156410150897	0.28109507549637	5.20741229198204
H	-0.84653549579726	-1.43742049504569	5.47639058867825
H	-0.77353544371805	-0.82440667801875	3.82624963644660
H	0.28996320343978	3.49325730249887	0.81656007950892
H	10.48822031141016	2.30379072600879	-0.38810337997132
H	14.58174998330308	0.05626815231622	1.29097580890792
H	5.68506847655677	3.50066790578062	-0.48280244060630

H	11.49269230607492	0.43574577196665	-1.64663890458743
H	13.50809550899110	-0.71426622736941	-0.79517626157628
H	0.96563070713007	4.95561231126557	-0.98458963550775
H	5.03422640747209	4.98022687341366	-2.30138698796600
H	2.70429343514398	5.78193685462903	-2.53828220947387
H	14.68940289088332	1.66375654734525	3.17840031992173

Table D.49. B3LYP/TZVP-optimized Cartesian coordinates of TOQ dimer in gas phase.

68
Coordinates from ORCA-job GeoOpt

O	13.24259674921726	5.69774798311154	5.24902695335252
O	11.15293315644463	6.51832501089984	5.57275522388372
N	10.27030972350127	4.81483182498700	3.53379701415139
N	10.17513502848561	4.00175467622097	2.49146024459490
N	11.42930975240564	3.53496595646666	2.13244012081363
C	11.53780986937860	4.91985836353235	3.86408799612393
C	12.32191367784807	4.09962009379981	3.00501405530252
C	13.69421600750858	3.80571647837686	2.90539684468767
C	14.10796446730506	2.97136433429544	1.91050658646503
C	13.19666877116313	2.46444949170972	0.93894484812735
C	11.81319988041202	2.79201594075597	1.00633269484960
C	11.93864164803234	5.81624692770268	5.00226386936785
C	13.77955193225089	6.51954626842397	6.35154651250225
C	15.25715339294228	6.23624936300030	6.46156852320490
C	8.89893829054837	3.38998982170472	2.22389089139448
C	7.82180228392537	4.24755255886086	2.01875542282312
C	6.54088740319103	3.72127506422431	1.92394777521655
C	6.36914540870925	2.34248007160899	1.99762559249049
C	7.44881941427173	1.48334091323150	2.18090238002967
C	8.72505751491990	2.01252197290235	2.31509151765790
C	13.64264338003251	1.68657355258727	-0.15164908983116
O	1.39027275192388	0.40227875865673	4.19610261749189
O	3.33042955348218	-0.28403747351047	5.14951818994641
N	4.70092883306577	1.10144420125719	3.14097017023637
N	5.05526771550708	1.75421356805736	2.04343210403486
N	3.93074801399510	2.14224773876095	1.33288421364261
C	3.38945960741446	1.02417362669028	3.17283707189082
C	2.84659413869827	1.69625950476774	2.04155659978926
C	1.54283844152982	1.94883280416783	1.57684157458551
C	1.39278570347486	2.62103532236165	0.40099254752517
C	2.51778677584291	2.99493888597678	-0.39029238697655

C	3.83927651970695	2.70414685940637	0.05120158249370
C	2.71419694567967	0.29605111839816	4.30119514770448
C	0.59424970778480	-0.26813001823477	5.24343684940929
C	-0.86431848114235	-0.04592448219160	4.92934601306269
C	10.93904663522120	2.42030696502023	-0.02157338771829
C	4.93925966898461	2.93320939124884	-0.78297260348298
C	11.42437235749704	1.67803266670040	-1.08051343475138
C	12.77230453101603	1.29180770966810	-1.14035593933753
C	2.35822686995041	3.59444428779015	-1.65837553601346
C	4.73312816684135	3.50138066700008	-2.02472252867664
C	3.44558959086170	3.85254702402579	-2.45924648784054
H	14.36558724098069	4.25208289126116	3.62250772576998
H	13.23217914216531	6.25075888266832	7.25501081884807
H	13.56799010473582	7.56210268561347	6.11465548036163
H	15.44639960787299	5.18608467923293	6.68807534556199
H	15.67160383790295	6.83379891035574	7.27547598154856
H	15.78605179155000	6.50799418258628	5.54650061802394
H	7.97702984671318	5.31765678875969	1.98856480960026
H	5.69543502531978	4.38743668716228	1.81572406367302
H	7.28593412327069	0.41795907828350	2.27390405635135
H	9.55967363849548	1.35258311764520	2.51082421659568
H	0.71240657176180	1.60028689056325	2.17101111350134
H	0.88761423387700	0.16191923614229	6.20090430520192
H	0.86927926749696	-1.32255986093432	5.23786847377036
H	-1.11647212458496	1.01553233695468	4.93124252985668
H	-1.46909507016363	-0.53412358461090	5.69552390322413
H	-1.13585159186935	-0.47825218613738	3.96505076554362
H	0.40196842372402	2.84804128332980	0.02779461767821
H	9.90463579012502	2.72257980973136	-0.00937107443261
H	14.69179133688860	1.42358572392588	-0.20490805166547
H	5.93523377373930	2.65408473783191	-0.48029298280731
H	10.75262421187307	1.39970476368688	-1.88243307864233
H	13.12797796154900	0.70297244762575	-1.97579542416683
H	1.35696353750417	3.82920353687626	-1.99708927924812
H	5.58107496023760	3.66834069394537	-2.67656816874885
H	3.30979715497897	4.30451627475480	-3.43284706042813
H	15.15513560602629	2.71027142032488	1.82080282080492

Table D.50. B3LYP/TZVP/RIJCOSX-optimized Cartesian coordinates of TOQ dimer in gas phase in the excited state.

68
Coordinates from ORCA-job PBE0

O	12.98924708372445	5.94423087732223	4.77144820528364
O	11.32681439175813	7.29418247114440	4.04298982580764
N	10.46828996795300	5.25081713691826	2.38350956276002
N	10.32379663502972	4.03254708545429	1.65845196235843
N	11.42395786912715	3.17843635993992	1.99232489186659
C	11.54157593080608	5.11405384237397	3.12448706878658
C	12.19199449276963	3.86591260073128	2.93385954054097
C	13.38424394796488	3.28047536989446	3.36855532714132
C	13.84122523437806	2.15544663383519	2.72761365495206
C	13.19920062648221	1.65643899643074	1.54738897259015
C	11.98781945296098	2.26085539061900	1.10329525657179
C	11.92629116287749	6.25417564309435	4.02423367011113
C	13.46372618571619	6.98961209784706	5.68694150088526
C	14.59236858483699	6.41178071423420	6.50550415651123
C	9.03519545921554	3.51527166622612	1.76005023618189
C	7.95149201925280	4.43761880765261	1.75451962962990
C	6.65636588912667	3.98663637447019	1.79607782626180
C	6.41196927472848	2.60385690891974	1.84815526178343
C	7.47173649467508	1.68183097592289	1.88924970085705
C	8.77266997714714	2.13385922604740	1.83582557582726
C	13.75704916683746	0.64387202027272	0.74940643048240
O	1.68925937262490	0.05525905168508	4.05192449489227
O	3.73986492643656	-0.67410875587390	4.69609662525314
N	4.87185060888457	1.17121549320780	2.93899181442187
N	5.10200580516381	2.08990464552266	2.01955359811675
N	3.91579983938132	2.54250769736001	1.47093786341757
C	3.56557801524574	1.02194384900492	3.05892326873258
C	2.90879906158090	1.90434723823967	2.15982958994172
C	1.56829925571918	2.18174789595238	1.88121025585287
C	1.27734453293888	3.07544540653905	0.87039127686854
C	2.29793394334277	3.62806428247048	0.06923706159409
C	3.67705380461566	3.32154079289432	0.31613370675667
C	3.01645172349103	0.02308692763523	4.04601400421714
C	1.01548205559007	-0.88890013517875	4.97867457039727
C	-0.47174400539274	-0.71216979156895	4.81444461884683
C	11.40540348295557	1.90114050400734	-0.10880522124103
C	4.65443000102849	3.72939742756519	-0.58227138451035
C	12.00168615106985	0.90768166944553	-0.87811406282694

C	13.16387068904786	0.26813147742721	-0.44335087599912
C	1.97624412592726	4.43196342640663	-1.05667255283587
C	4.28601664103562	4.48539538972082	-1.69282127883629
C	2.95519456802263	4.85952349771130	-1.92038285263006
H	13.93749174609307	3.75421328746579	4.16448546074626
H	12.62086167953168	7.30222231096888	6.30283594383629
H	13.77951088076297	7.84002189642643	5.08190814913274
H	14.25718277751580	5.55784043758491	7.09620368541784
H	14.95847073425436	7.17477689398097	7.19497587296019
H	15.42650002605499	6.10274576426366	5.87417563817025
H	8.16994230790064	5.49513407044860	1.75774673972066
H	5.84586157337180	4.70225848382744	1.83682644614463
H	7.26541504955932	0.62311423347432	1.96951520380747
H	9.58375175337729	1.42366942467386	1.86692941052485
H	0.80779117505672	1.68105050569661	2.46180625576989
H	1.36210071793224	-0.65239482438469	5.98422051940645
H	1.35421575555601	-1.89093631680168	4.71729991173530
H	-0.78632647476197	0.30011795622762	5.07208057295679
H	-0.98218187476833	-1.40033584863251	5.48998743185929
H	-0.79320734093665	-0.94583055890648	3.79840053557313
H	0.24635223745406	3.32559350298249	0.65310738858666
H	10.52377634097948	2.41414174089070	-0.46656847070408
H	14.67885497617470	0.17655191609613	1.07198800496645
H	5.69156727756823	3.47294348058505	-0.44440528868000
H	11.56300456360686	0.63791121759267	-1.82978559714812
H	13.61318460402791	-0.50672909615129	-1.05007340287502
H	0.93634986161171	4.67582714040939	-1.23456992890525
H	5.05415532560074	4.78592995273692	-2.39377858169954
H	2.70175438786694	5.46022127118460	-2.78366983111261
H	14.75987569453132	1.68515918983454	3.05263003218923

CHAPTER E. APPENDIX FOR CHAPTER 6

TTTTGTTTANCTTTAAGA*AGGAGA*TATACC**ATG**GGTCATCATCATCATCATCA

CGATTACGATATCCCAACGACCGAAAACCTGTATTTTCAGGGCGCCCATATG

GGAATTCAAAGGCCTACGTCGACC**GTTGAAAACAATAAATGCCGTGAAGA**

TTTCCGCTTTACGCAAGAATACGAAGAAGACTACCCGAACACCAATGAA

CGCTACTACGAAAACTATCAGGTGGCCGATCGTTACTACAACTACCCGA

ACAAATACAAAGAACCGAAAATCAAACAATGCTGTTGCAAAAAATCCAT

GCGCGAAGCTCTGGAACTGCTGCGTTACGATGCGCTGCGCCCGTTTGTT

AACTTCAATCAGTTTGCGTTCATCAGCGATTTCTTTATTGTGGGCGCCAA

CCTGGTTGGTATCGACCTGAGCGCACCGCCGAAAGATAATCTGTCTGGC

CTGGACGGCACCTTTGAACGTTTCAGCGCTTGCAACTGTGATCTGATTG

ACATCGCAGGCCGCGTCTCTTATCCGATTCCGGTGCCGCTGACGCTGGA

AGGTCTGATTAATACCATCGGCACGATTCCGGGTGTGGCAGAACTGATC

GCTCTGATTGATGCGGTTATCCCGCCGACCATTGATCTGGGCGCCATCC

TGGACGCAATTCTGGCGGCCATTATCGACTTTATCCTGGCAGCTAGTAC

CCCGCTGGCTAACGTCGATCTGGCGTCCCTGTGCAATCTGAAAGCGGTT

GCCTTTGACATCACGCCGGCCGATTATGAAGACTTCATTGCAAGTCTGG

GTTACTACCTGGATAAAAAACATTACAAAGAATGCAACTGTAATTGCGAC

TGTGATGACTGTTGCTGTAACAAAGGCATCCTGGATAACCTGTACATGTC

AAACATCAACAACCAAGTTACCGTGGTTGCCGGTTCGCTGGTCCTGACG

GGCGTCGAAGTGCTGGGCAAGAAAAACGATGTTATTGTCCTGGGTAACT

CCAATGACAGCCGCATCTATTTCGTGTGTGTTGACTCCATCGACTACATC

GCCTAAGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGC

TGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTC

TTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATATCCCGCAAGAG

GCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAGGGTGACG

GTGCCGAGGATGACGATGAGCGCATTGTTAGATTTCATACACGGTGCCTGAC

TGCGTTAGCAATTTAACTGTGATAAACTACCGCATTAAAGCTTATCGATGATA

AGCTGTCAAACATGAGAATTCTTGAAGACGAAAGGGCCTCGNGATACGCCTA

TTTTTATAG

Italics: Ribosomal binding site; **ATG:** start codon; Underlined: His6-tag; GTCGAC: SalI site; **Bold**: CotA sequence; GGATCC: BamHI site

Figure E.1. CotA sequence in the pET15b plasmid.

CHAPTER F. APPENDIX FOR CHAPTER 7

Start S Y Y H H H H H H L E S T S L Y K K A G S A A A V L E E N L Y F Q G S F T
D E P P A C L T S T A D L S V D I F T D K V E P L F S Q G W N V T Y H N T Y K I A N
N L F D N T T Y L L Y Q C G S T P P A D V V D N G N F N A V L E I P L S N V G L S Q
T P H I G F M E Q L E L V D E I A A F L T D T D F I S S P C F L D E I A A G N V L T L
V E P S E G V D A P A T G N T A L S A G T V A F V A S F T Q V P F D N T V N I Q E Y
S E L T N V A V F E W V K F F S L F F N K E H T A N Q V V E A A E S R F D C V A Q
N A G A V Q A D N M P V Q P V V L W A Y Y S D F C G G W D V A E C P N Y Y C E F
A N A C G A E I I S S T E G N T T V C G A P Y M T T E E L V E L G K D A D H W I Y P
S N N W D T A S E T F G E Q L Q N M K A V Q D Q Q V F D Y Q A S G E N A W F E Q
R Y A E Y Y N V L A D F C A V V G T T Q P L T G R S W F R N V F T E P V G S L P D
C S P T Q S A N I L D D V H I C F L P T T G G A A A G G G S G S G G S S A K A I A V
G T A A L A A G L L S L I H V L L F *Stop*

Figure F.1. Amino acid sequence of CBA1. Blue text indicates a short linker after the start codon, green text indicates His_6-tag, orange text indicates linker, yellow text indicates TEV cleavage site, red text indicates a short linker. The CBA1 sequence is in black text.

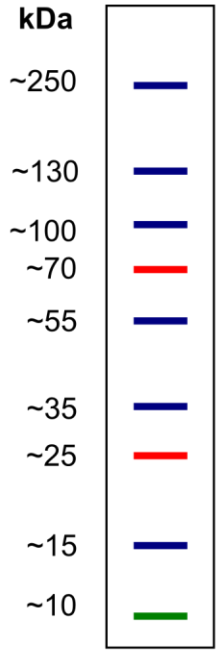


Figure F.2. PageRule Plus Prestained Protein Ladder (ThermoFisher) used for SDS-PAGE. Molecular weight range displayed from 10 to 250 kDa.

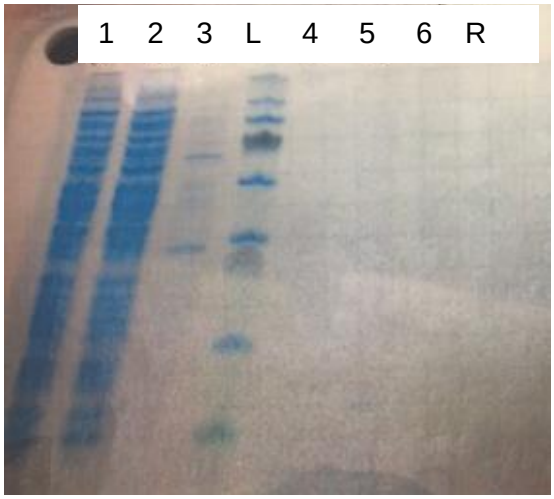


Figure F.3. SDS-PAGE gel of CBA1 expressed through IPTG-induction and lysed via sonication. Purification used Ni(II)-NTA affinity column. Equilibration buffer was 100 mM NaCl, 100 mM Tris at pH 7.4. Lane titles signify the following: 1=filtrate; 2= flow-through; 3= 50 mM imidazole wash; L= molecular weight ladder; 4= 100 mM imidazole wash; 5= 250 mM imidazole wash; 6= 400 mM imidazole wash; R= regeneration wash.

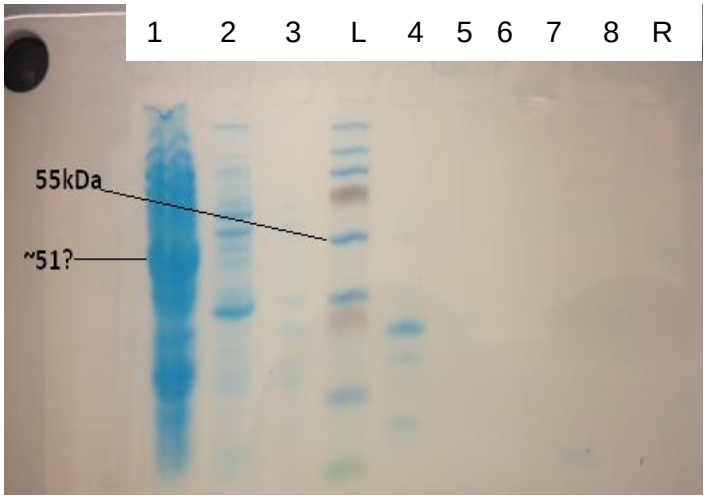


Figure F.4. SDS-PAGE gel of CBA1 expressed through autoinduction and lysed via sonication. Purification used Ni(II)-NTA affinity column. Equilibration buffer was 100 mM NaCl, 100 mM Tris at pH 7.4. Lane titles signify the following: 1= lysate; 2= 10 mM imidazole wash; 3= 25 mM imidazole wash; L= molecular weight ladder; 4= 50 mM imidazole wash; 5= 75 mM imidazole wash; 6= 100 mM imidazole wash; 7 = 250 mM imidazole wash; 8 = 400 mM imidazole wash; R = regeneration wash.

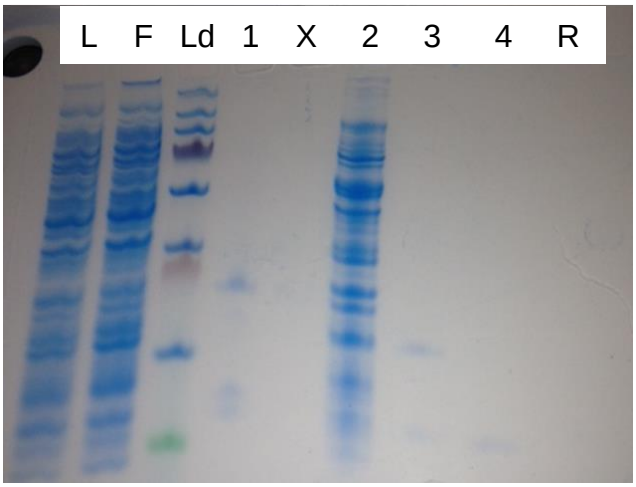


Figure F.5. SDS-PAGE gel of CBA1 expressed through L-arabinose-induction and lysed via sonication. Purification used a Ni(II)-NTA affinity column. Equilibration buffer was 100 mM NaCl, 100 mM Tris at pH 7.4. Lane titles signify the following: L= lysate; F=filtrate; Ld= molecular weight ladder; 1= 100 mM imidazole wash; X= empty lane; 2= 50 mM imidazole wash; 3= 250 mM imidazole wash; 4= 400 mM imidazole wash; R= regeneration wash.

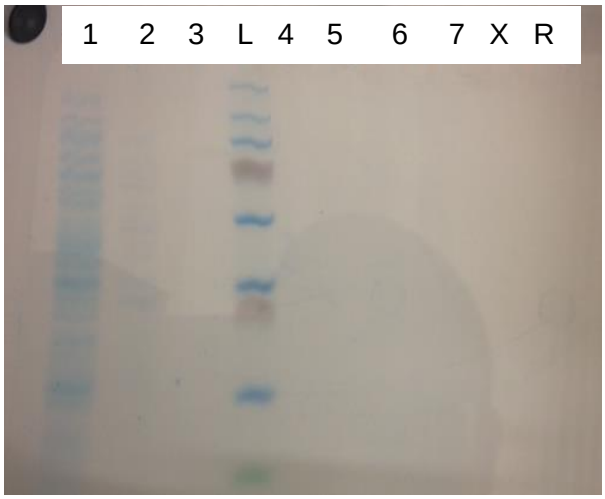


Figure F.6. SDS-PAGE gel of CBA1 expressed through autoinduction and lysed enzymatically. Purification used Ni(II)-NTA affinity column. Equilibration buffer was 100 mM NaCl, 100 mM Tris at pH 7.4. Lane titles signify the following: 1=lysate; 2= 10 mM imidazole wash; 3= 25 mM imidazole wash; L= molecular weight ladder; 4= 50 mM imidazole wash; 5= 75 mM imidazole wash; 6= 100 mM imidazole wash; 7 = 400 mM imidazole wash; X= empty well; R = regeneration wash.

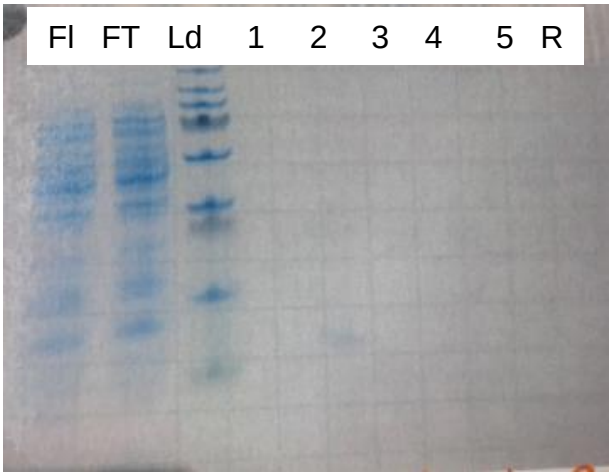


Figure F.7. SDS-PAGE gel of CBA1 expressed through L-arabinose-induction and lysed via sonication. Purification used Co(II)-NTA affinity column. Equilibration buffer was 400 mM NaCl, 110 mM KCl, 10 mM imidazole, 10 % (*v/v*) glycerol, 0.6 % (*v/v*) Triton X-100, 50 mM potassium phosphate buffer at pH 7.8. Lane titles signify the following: Fl=filtrate; FT= flow-through; Ld= molecular weight ladder; 1= 25 mM imidazole wash; 2= 50 mM imidazole wash; 3= 100 mM imidazole wash; 4= 250 mM imidazole wash; 5= 400 mM imidazole wash; R= regeneration wash.

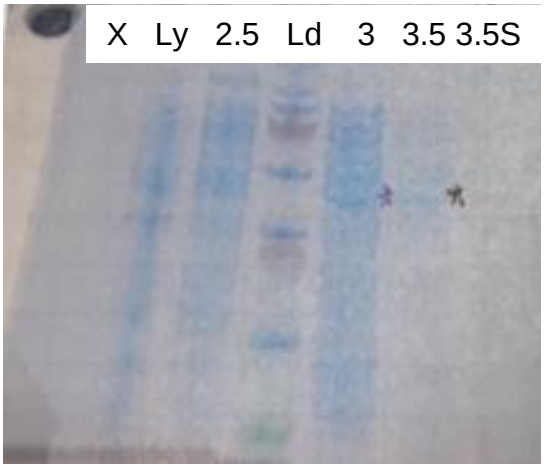


Figure F.8. SDS-PAGE gel of CBA1 expressed through L-arabinose-induction and lysed via sonication and enzymatically. An ammonium sulfate gradient was used to induce "salting-out". Purification used Co(II)-NTA affinity column. Equilibration buffer was 400 mM NaCl, 110 mM KCl, 10 mM imidazole, 10 % (*v/v*) glycerol, 0.6 % (*v/v*) Triton X-100, 50 mM potassium phosphate buffer at pH 7.8. Lane titles signify the following: X=empty well; Ly= lysate with 0 M $AmSO_4$; 2.5= pelleted lysate with 2.5 M $AmSO_4$; Ld= molecular weight ladder; 3= pelleted lysate with 3 M $AmSO_4$; 3.5= pelleted lysate with 3.5 M $AmSO_4$; 3.5S= supernatant of lysate with 3.5 M $AmSO_4$. Pelleted lysates were resuspended in 50 mM Tris buffer at pH 7.4.

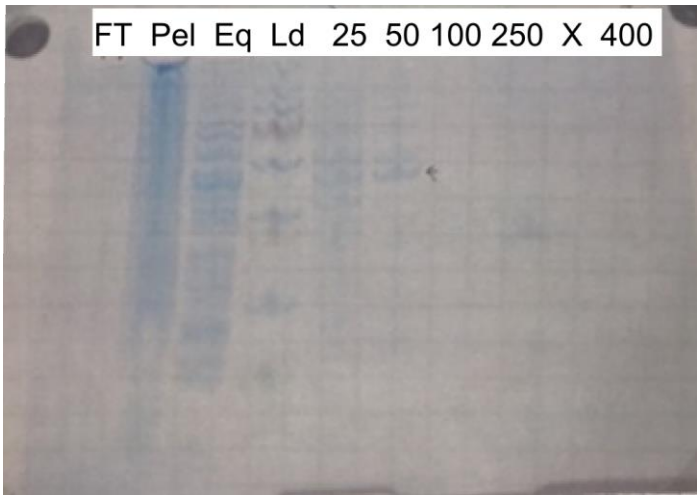


Figure F.9. SDS-PAGE gel of CBA1 expressed through L-arabinose-induction and lysed via sonication and enzymatically. An ammonium sulfate concentration of 2.5 M was used to induce "salting-out". Purification used Co(II)-NTA affinity column. Equilibration buffer was with 50 mM Tris buffer at pH 7.4. Lane titles signify the following: FT= column flow-through; Pel= resuspended pellet with 50 mM Tris buffer at pH 7.4; Eq= 0 mM imidazole wash; Ld = molecular weight ladder; 25= 25 mM imidazole wash; 50 = 50 mM imidazole wash; 100= 100 mM imidazole wash; 250 = 250 mM imidazole wash; X = empty well; 400= 400 mM imidazole wash.

Table F.1. Recipes for lysis buffers used to resuspend CBA1 cell pellet.

Buffer	NaCl	Tris	Detergent	KH_2PO_4	K_2HPO_4	Glycerol	KCl	Imidazole
1	350 mM	100 mM	-	-	-	-	-	-
2	350 mM	100 mM	0.1 % (*v/v*) Triton-X	-	-	-	-	-
3 pH 7.8	390 mM	-	0.6 % (*v/v*) Triton-X	3.6 mM	47 mM	10 % (*v/v*)	100 mM	10 mM

Table F.2. Recipes for buffers to equilibrate Ni(II)-NTA affinity column. Equilibration buffers used as initial recipe for equilibration buffers.

Buffer	NaCl	Tris	Detergent	KH_2PO_4	K_2HPO_4	Gly	KCl
1	100 mM	100 mM	-	-	-	-	-
2 pH 7.8	440 mM	-	0.6 % (*v/v*) Triton-X	3.6 mM	47 mM	10 % (*v/v*)	110 mM

Table F.3. Purification washes used to separate His_6-tagged CBA1 from native proteins expressed by *E. coli*.

Buffer	NaCl	Tris	Imidazole	Detergent	KH_2PO_4	K_2HPO_4	Gly	KCl
1	100 mM	100 mM	10 mM 25 mM 50 mM 75 mM 100 mM 400 mM 500 mM	-	-	-	-	-
2	100 mM	100 mM	10 mM 25 mM 50 mM 75 mM 100 mM 400 mM	-	-	-	-	-
3	100 mM	100 mM	50 mM 100 mM 400 mM	-	-	-	-	-
4	100 mM	100 mM	50 mM 100 mM 250 mM 400 mM	-	-	-	-	-
5	100 mM	100 mM	50 mM 100 mM 250 mM 400 mM	-	-	-	-	-
6	440 mM	-	0 mM 50 mM	0.6 % (*v/v*) Triton-X	3.6 mM	44 mM	10 % (*v/v*)	110 mM
7	440 mM	-	0 mM 50 mM 100 mM 400 mM	0.6 % (*v/v*) Triton-X	3.6 mM	44 mM	10 % (*v/v*)	110 mM
8	440 mM	-	0 mM 25 mM 50 mM 100 mM 250 mM 400 mM	0.6 % (*v/v*) Triton-X	3.6 mM	44 mM	10 % (*v/v*)	110 mM

Table F.4. Total protein precipitated out during ammonium sulfate salting-out procedure. Pellet was resuspended in 50 mM Tris buffer at pH 7.4. Concentration determined using a NanoDrop Spectrophotometer.

Concentration of $AmSO_4$ (M)	Total protein in resuspended pellet (mg/mL)
0 (lysate)	10.8
1.5	*no pellet*
2.0	*no pellet*
2.5	3.3
3.0	12.0
3.5	6.1
4.0	2.9

CHAPTER G. CHAPTER 8 APPENDIX

SECTION G.1 METHODS

Preparing the NMR spectrometer to detect cobalt-59 nuclei. The Bruker Avance 500 NMR spectrometer at the University of Vermont (UVM) was configured to detect the ^{59}Co nucleus using the cobalt-59 standard. The 90° pulse was optimized using the standard consistent with the literature, $K_3[Co(CN)_6]$ in D_2O.

Referencing the ^{59}Co Spectra. Previous publications indirectly referenced the sample of interest to an external sample 1 M $K_3[Co(CN)_6]$.[1] There are many disadvantages to this external referencing methodology as it can introduce error to the spectra and may inadvertently alter some acquisition parameters. Internal referencing of the Co(III) tetrapyrroles was attempted by including a single crystal of the standard along with the species of interest. This was unsuccessful as the strength of the radiofrequency pulse was insufficient to excite both the tetrapyrrole of interest and the standard.

IUPAC proposed a unified scale to report chemical shifts for nuclei other than ^{1}H in hopes to implement a consistent and accurate referencing methodology to increase the ease and effectiveness of communication across scientific disciplines.[2] This scale is formally based on the ratio of the X-isotope frequency to that of ^{1}H of tetramethylsilane (TMS) in $CDCl_3$ in the same magnetic field.[3] Sodium-3-(trimethylsilyl) propanesulfonate (DSS; $\delta = 0.00$) is a suitable substitute for aqueous solutions.[3] DSS has been known to interfere with biological samples such as proteins, in which case it is acceptable to have DSS as an external reference.[4] Both of these inert compounds are excellent NMR standards as their chemical shifts are impervious to variations in temperature and pH. X-nuclei are

referenced to the ^{1}H of DSS through the recalculation of the frequency which is termed the Ξ -value. Ξ is often expressed as a percentage (**Eqn. G.1**).

$$\Xi(\%) = 100\left(\upsilon_{obs}^{X} \Big/ \upsilon_{obs}^{DSS}\right) \qquad \textbf{Equation G.1}$$

The basis of Ξ referencing is to determine the zero frequency of the ^{59}Co chemical shift by multiplying the experimentally determined 1H frequency by the ^{59}Co to ^{1}H ratio; the Ξ ratio for ^{59}Co to ^{1}H is 23.727074 %.[3] The recalculation of the absolute frequency for the heteronuclei is show in **Equation G.2**.

$$\upsilon_{0}^{X} = \Xi\upsilon_{0}^{H} \qquad \textbf{Equation G.2}$$

The simplest reference technique is internally referencing to the solvent. The water peak in the ^{1}H spectra was calibrated to match the definite shift recorded from the ^{1}H of DSS which was dissolved in Tris-base to ensure the same buffer sample conditions. All samples were scanned at 25 °C and at pH 7.5. The water peak was then used as the ^{1}H frequency to which the referenced frequency of cobalt-59 spectra was recalculated. If the absolute frequency of DSS was measured at 500.1366624 MHz at 0.0 ppm, the absolute zero frequency of ^{59}Co would be 118.667796 MHz as determined by **Equation G.2**. Coupling Ξ with solvent referencing is not only accurate, convenient, and time efficient, but also offers a consistent methodology to span across the disciplines.

SECTION G.2. ABS SPECTRA OF COBALAMINS

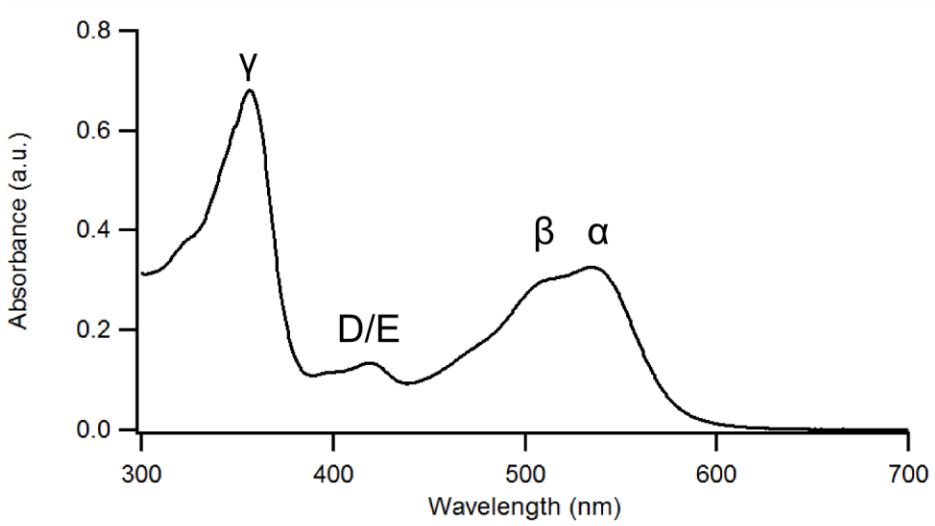


Figure G.1. Typical Abs spectrum of cobalamin.

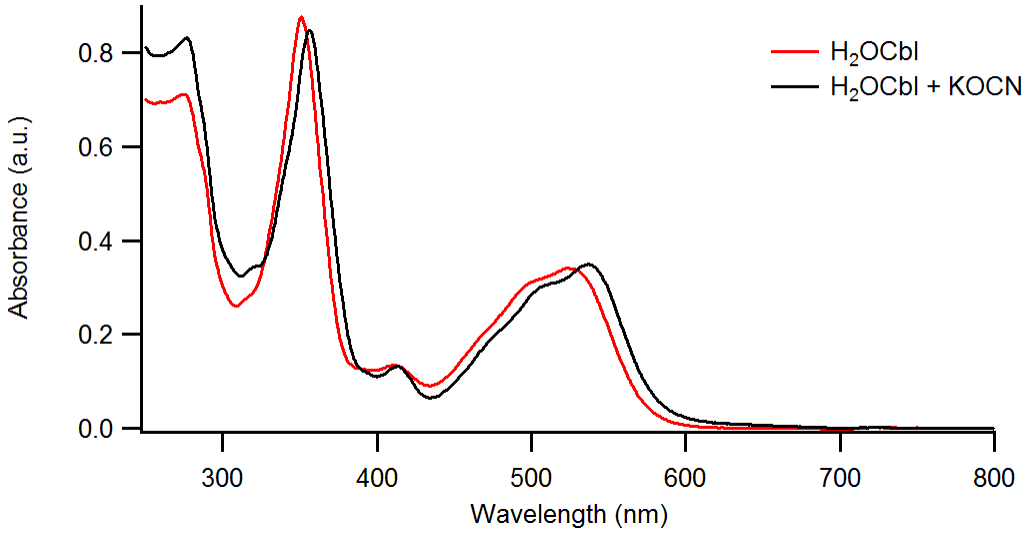


Figure G.2. Abs spectra of H_2OCbl and KOCN + H_2OCbl in sodium phosphate buffer at pH 8.0. Maxima of Abs spectra of H_2OCbl at 351, 410, and 523.5 nm. Maxima of Abs spectra of 20 KOCN: 1H_2OCbl at 357, 413, and 537 nm.

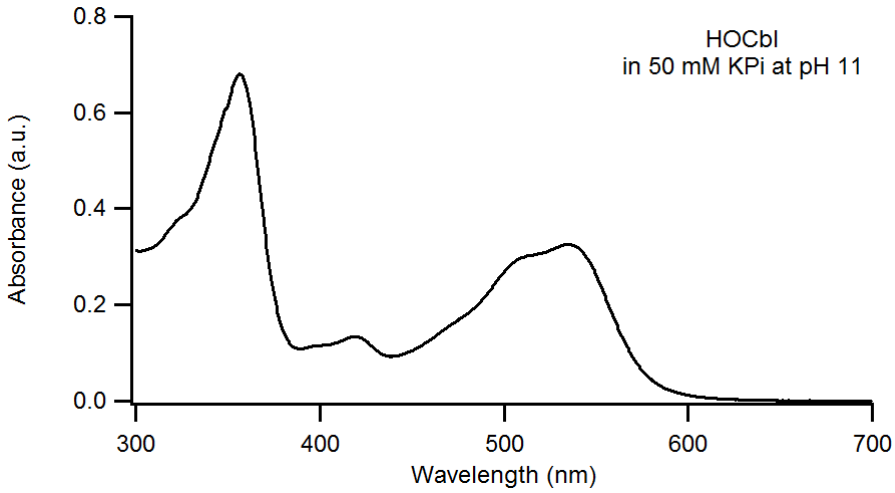


Figure G.3. Abs spectra of HOCbl in potassium phosphate (KPi) buffer at pH 11. Maxima of Abs spectra of HOCbl at 356.5, 419.5, and 535 nm.

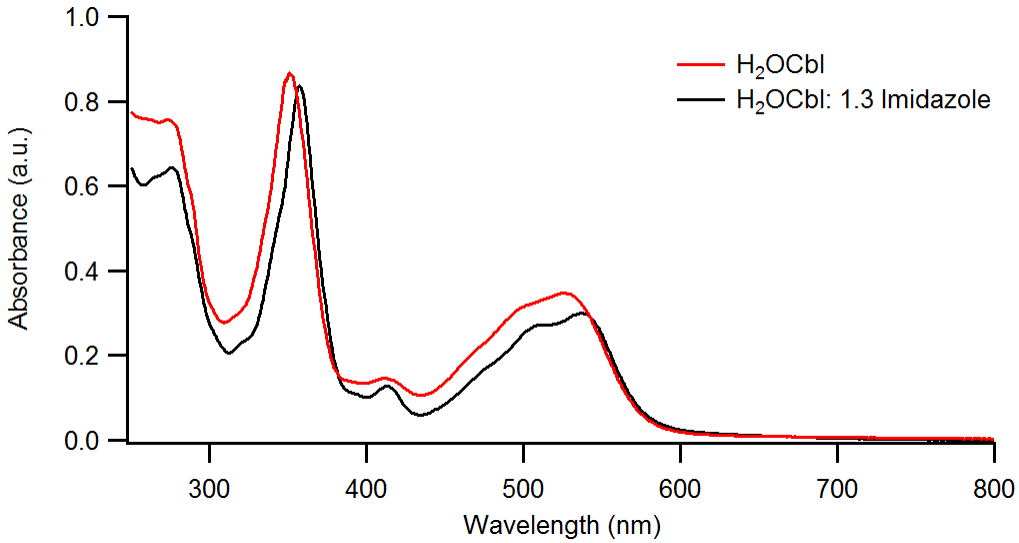


Figure G.4. Abs spectra of H_2OCbl and imidazole + H_2OCbl in Tris buffer at pH 7.5. Maxima of Abs spectra of H_2OCbl at 351, 411, and 526 nm. Maxima of Abs spectra of 1.3 imidazole: $1H_2OCbl$ at 357, 413, and 538 nm.

Table G.1. Published maxima of cobalamins in water unless otherwise noted.

Cobalamin	Peak maxima (nm)	Solvent	Reference
OCNCbl	272-278 353 520-530	Water	Ref[5]
HOCbl	279 325 359 421 516 537	Water	Ref[6]
H_2OCbl	274 317 351 411 499 527	Water	Ref[6]
ImCbl	357 413 536	Tris/HCl at pH 7.5 with 1 M KCl	Ref[7]
bisImCbl	360 536	Tris/HCl at pH 7.5 with 1 M KCl	Ref[8]
CNCbl	277 360.5 520 552	Water	Ref[9]
diCNCbl	367 415 540 579	Water	Ref[10]
Co(I)Cbl	386.5 455 545 680 800	60:40 mixture of glycerol and NH_4Cl	Ref[11]

SECTION G.3. ^{1}H NMR SPECTRA

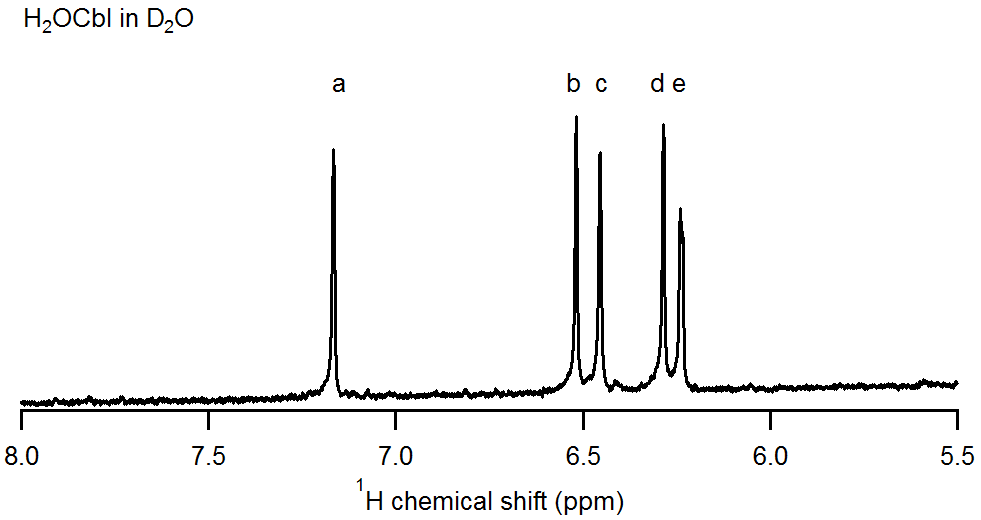


Figure G.5. Aromatic region of ^{1}H NMR spectra of H_2OCbl in D_2O. The chemical shifts for the labeled aromatic peaks of H_2OCbl are: (a) 7.17, B2; (b) 6.52, B7; (c) 6.45, B4; (d) 6.29, C10; (e) 6.24d, R1.

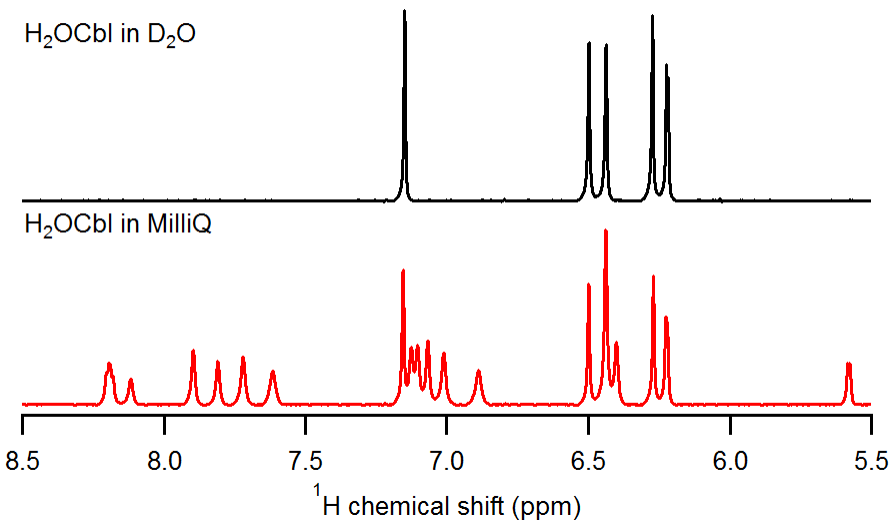


Figure G.6. Aromatic region of ^{1}H NMR spectra of H_2OCbl in MilliQ water and D_2O. The chemical shifts for the labeled aromatic peaks of H_2OCbl are: (a) 7.17, B2; (b) 6.52, B7; (c) 6.45, B4; (d) 6.29, C10; (e) 6.24d, R1.

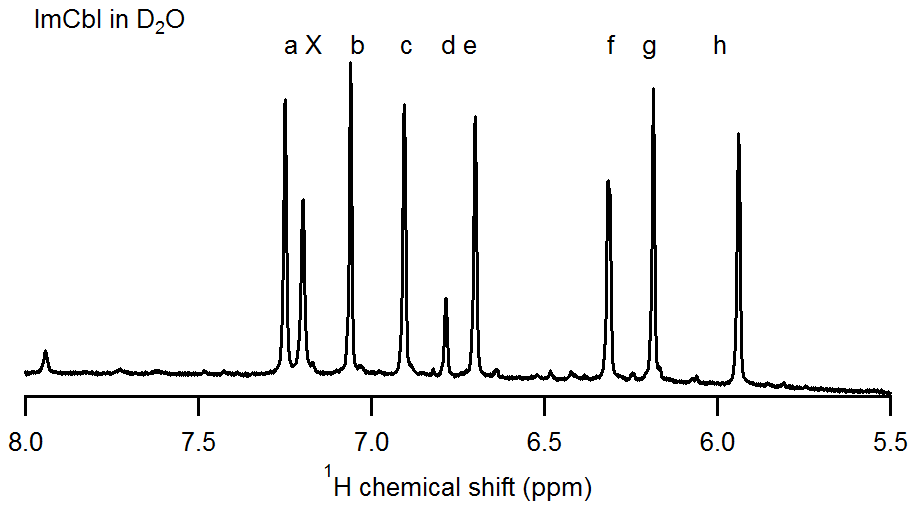


Figure G.7. Aromatic region of ^{1}H NMR spectra of ImCbl in D_2O. The chemical shifts for the labeled aromatic peaks of ImCbl are (a) 7.25; (X) 7.20, free imidazole (b) 7.06, B2; (c) 6.91, bound imidazole; (d) 6.78, bound imidazole; (e) 6.70, B4; (f) 6.32d, R1; (g) 6.18, C10; (h) 5.94, bound imidazole.

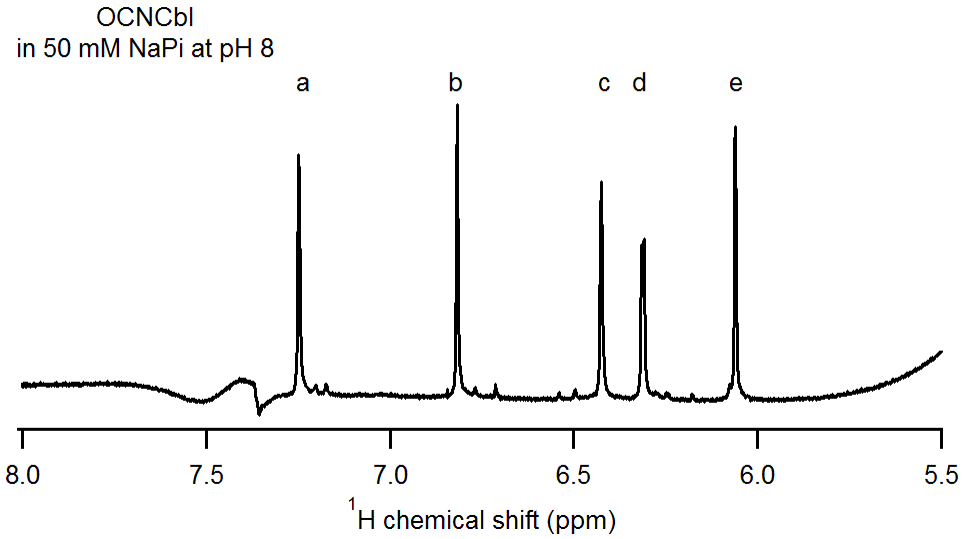


Figure G.8. Aromatic region of ^{1}H NMR spectra of OCNCbl in 50 mM sodium phosphate buffer at pH 8.0 (10 % D_2O). The chemical shifts for the labeled aromatic peaks of OCNCbl are (a) 7.25; (b) 6.82; (c) 6.43; (d) 6.31d; (e) 6.06. Impurities located at 7.20, 7.17, 6.84, 6.77, 6.71, 6.54, 6.50, 6.28, 6.25d, 6.18, and 6.08 ppm. Some of these impurities can be assigned to the starting material.

SECTION G.4. DFT CALCULATIONS

Table G.2. DFT predicted bond lengths of the simplified cobalamin structure. Bond number 1 and 2 are the upper and lower axial ligands respectively. Bonds 3-6 are the equatorial bonds.

Bond #	CNCbl (Å)	diCNCbl (Å)	MeCbl (Å)	ImCbl (Å)	H_2OCbl (Å)
1	1.85	1.94	1.98	2.00	2.13
2	2.14	1.94	2.27	2.01	1.94
3	1.90	1.95	1.89	1.91	1.90
4	1.94	1.94	1.94	1.95	1.95
5	1.94	1.90	1.93	1.95	1.95
6	1.90	1.90	1.89	1.90	1.90

Table G.3. The cobalt bond lengths from X-ray crystal structures of CNCbl, ImCbl and MeCbl. No published crystal structures are available for diCNCbl or H_2OCbl.

Bond #	CNCbl (Å)	MeCbl (Å)	ImCbl (Å)
1	1.81	1.98	1.94
2	1.98	2.16	2.01
3	1.89	1.87	1.89
4	1.91	1.92	1.97
5	1.91	1.92	1.94
6	1.89	1.88	1.91
ImCbl: Ref[7] MeCbl: Ref[12] CNCbl:Ref[13]			

Table G.4. PBE/TZVP-optimized Cartesian coordinates of cyanocobalamin.

Co	4.571493003	2.187536599	5.229653623
C	4.113099787	2.712849718	3.512311056
N	3.817465574	3.061547544	2.432195872
C	3.000985952	4.272439011	6.505469830
C	1.520996394	4.797221742	6.307243940
C	0.703209984	3.459368169	6.278638540
C	1.715438601	2.463553356	5.740506370
C	1.361601668	1.201511772	5.141952682
C	2.323898826	0.280244764	4.764622176
C	2.073279442	-1.080861141	4.084910558
C	3.453645686	-1.778200411	4.285340307
C	4.369248102	-0.592456599	4.418938364
C	5.719412061	-0.641931127	4.091002708
C	6.608079472	0.424615034	4.106647880
C	7.993718327	0.376661495	3.480940639
C	8.662718411	1.595558819	4.178998391
C	7.463561305	2.452271079	4.567913767
C	7.537928947	3.793380515	4.905690393
C	6.260671395	6.059480816	5.483078005
C	4.866159400	6.190091843	6.189798871
C	4.116767910	4.956640751	5.677932172
C	3.409465557	4.183709082	7.983381660
N	2.941493815	2.899806697	5.900863910
N	3.694856024	0.494643709	4.851696510
N	6.332950758	1.652221410	4.602920168
N	5.199492087	3.958538822	5.506958771
C	1.031169718	5.759242112	7.394551763
C	1.319738736	5.465847126	4.931669968
C	0.059654633	2.978150939	7.591311062
C	-0.109473434	0.932265081	4.898410205
C	0.917978468	-1.937991981	4.618059355
C	1.888102451	-0.819288059	2.569286356
C	3.588241306	-2.685653758	5.522166506
C	8.722418362	-0.958441621	3.661579243
C	9.499091902	1.242706441	5.420900302
C	8.871438905	4.503690184	4.807178646
C	6.254797294	6.698929569	4.076639892
C	7.331071760	6.708541844	6.374472193
C	4.170022186	7.533292169	5.994140405
C	6.378604635	1.749265634	7.651885430
N	5.167441563	1.505115111	7.170497988
C	3.214098564	0.231718127	8.228199147
C	2.824952348	-0.486196847	9.361267225

C	3.737168656	-0.694001922	10.443799510
C	5.029734817	-0.170399404	10.380851192
C	5.395874626	0.546798722	9.240631581
C	4.514128546	0.751104866	8.154415658
C	1.427115450	-1.044464065	9.443283766
C	3.308555761	-1.474182942	11.657966775
C	6.366825290	4.528667459	5.316330280
N	6.568297912	1.197752670	8.883703938
C	7.834400692	0.677899679	1.970281535
H	-0.110808465	3.590672813	5.550101388
H	3.709483772	-2.365366895	3.389538433
H	6.089399465	-1.590516478	3.702868887
H	9.302294126	2.123245215	3.456445326
H	5.058518871	6.047881418	7.266644585
H	3.718349796	5.150364188	4.669402925
H	2.749165837	3.510629919	8.540972288
H	3.362831904	5.172837574	8.457012639
H	4.434355509	3.805988152	8.079329207
H	1.615752607	6.689132533	7.402544432
H	1.071003533	5.326804451	8.401768142
H	-0.013530000	6.042307639	7.195843957
H	1.869643288	6.412582215	4.850049573
H	1.622999093	4.815442572	4.096877723
H	0.799500859	2.777427975	8.377135303
H	-0.652233715	3.724021986	7.969870378
H	-0.284335113	0.494044794	3.908727861
H	-0.690547116	1.859038820	4.928104078
H	-0.550382258	0.251482388	5.642853460
H	0.904681043	-1.992654463	5.715321507
H	1.026618458	-2.962995811	4.233323521
H	-0.058320750	-1.577639018	4.279568678
H	2.741532339	-0.268849812	2.148374461
H	1.794589863	-1.779148295	2.039811403
H	3.359824563	-2.141786994	6.449974864
H	2.920858580	-3.554753008	5.450459969
H	8.766621192	-1.278080504	4.711548849
H	9.750674681	-0.885163609	3.279526910
H	8.226429167	-1.753205931	3.086871959
H	8.901146327	0.682849747	6.156232814
H	9.878597660	2.150787350	5.908974341
H	9.290180929	4.746510691	5.796182158
H	8.791316861	5.448246686	4.254484200
H	9.613256278	3.895669344	4.280204643
H	7.200505216	6.513848318	3.548918180

H	6.133773047	7.788013962	4.158492132
H	5.445424992	6.304869962	3.445127674
H	8.297234135	6.829281917	5.872269569
H	6.993844236	7.714830449	6.664026316
H	3.859005928	7.691324464	4.952303603
H	4.834071917	8.364345777	6.275061089
H	7.148503813	2.319052826	7.141468764
H	2.508943372	0.386257871	7.412156110
H	5.726487828	-0.319573208	11.208446156
H	0.885330298	-0.657564586	10.320774892
H	1.432323099	-2.142204658	9.536120342
H	0.844239175	-0.786186829	8.548708321
H	4.123218639	-1.552222344	12.389924783
H	2.997043397	-2.496714627	11.392634571
H	2.448768282	-1.002395567	12.159413468
H	0.253279918	5.700528635	4.797734782
H	-0.502072168	2.048111586	7.424097897
H	0.983237063	-0.230861929	2.366542863
H	4.617383450	-3.062634979	5.600097552
H	10.365804082	0.624649000	5.150946930
H	7.487972092	6.133530989	7.299231752
H	3.276424191	7.613147317	6.626046997
H	7.418177325	1.260765337	9.435535380
H	7.329967673	1.640492211	1.800207767
H	7.242557180	-0.104365503	1.475015019
H	8.825117613	0.712296716	1.492461315

Table G.5. PBE/TZVP-optimized Cartesian coordinates of dicyanocobalamin.

Co	0.000000000	0.000000000	0.000000000
C	-0.000000000	0.000000000	-1.940634298
N	-0.006091344	-0.036619144	-3.114011366
C	0.691696679	-2.800029325	0.219195981
C	1.798202953	-3.710717818	-0.446455807
C	3.090966944	-2.841423083	-0.251348088
C	2.532854184	-1.426473741	-0.225940880
C	3.322419047	-0.241800057	-0.451347932
C	2.785182398	1.026494239	-0.313163942
C	3.528063493	2.361550665	-0.537675674
C	2.541258861	3.370513257	0.123333923
C	1.237840273	2.622373613	0.031958181
C	0.000000000	3.256323387	0.071943382
C	-1.242475765	2.634306018	0.071267058

C	-2.568976407	3.381125623	0.050115384
C	-3.504914864	2.287838167	0.639709161
C	-2.772433764	1.003939970	0.271983925
C	-3.356749231	-0.248879206	0.238902124
C	-3.110939898	-2.844663478	-0.322191599
C	-1.808717707	-3.706959575	-0.178568902
C	-0.689801067	-2.705525273	-0.486774031
C	0.507118246	-3.078364708	1.719558663
N	1.250742963	-1.431072867	0.020822224
N	1.451202247	1.297365606	-0.060322756
N	-1.439382175	1.301976678	0.068944046
N	-1.266723974	-1.412013675	-0.079131382
C	1.925008867	-5.107605964	0.169169325
C	1.587608035	-3.858107636	-1.967031571
C	3.980655173	-3.142340468	0.967501042
C	4.764383150	-0.439663440	-0.871364071
C	4.941524355	2.481017775	0.048834110
C	3.550331987	2.644355404	-2.059387858
C	2.817276887	3.720394127	1.597574692
C	-2.559905542	4.704294739	0.819732877
C	-3.677883634	2.349669918	2.167896802
C	-4.846640411	-0.375537022	0.469574631
C	-3.643709852	-2.832948279	-1.770684193
C	-4.171562130	-3.396053005	0.644258076
C	-1.795144432	-4.994339283	-0.998304106
C	-0.007205314	0.100556662	1.933268112
C	-2.571452938	-1.433848383	-0.004370590
C	-2.949562086	3.642935806	-1.425899622
N	-0.028791678	0.238019729	3.098882913
H	3.715191398	-2.974915935	-1.148680956
H	2.502498082	4.301230478	-0.465318464
H	0.012685239	4.345235749	0.110853021
H	-4.491023374	2.349409491	0.156559416
H	-1.743485996	-3.969985469	0.889777673
H	-0.513729741	-2.641265931	-1.573545199
H	1.448085038	-2.937482303	2.262658050
H	0.168775876	-4.110406910	1.884300956
H	-0.222412811	-2.385492023	2.155870652
H	0.998668823	-5.685111653	0.040539789
H	2.157972334	-5.082182196	1.240769396
H	2.726938613	-5.668822793	-0.335469448
H	0.692085211	-4.449160305	-2.202401470
H	1.501061228	-2.884860733	-2.473855695
H	3.455952051	-2.990251015	1.920088048

H	4.349255748	-4.177502445	0.936619163
H	5.072018053	0.286721414	-1.632987013
H	4.910826499	-1.429150708	-1.318714898
H	5.473961678	-0.357606291	-0.032117431
H	4.995433771	2.105860481	1.080120547
H	5.240635201	3.540511880	0.059002033
H	5.686962347	1.946416280	-0.548874212
H	2.534965111	2.628160197	-2.480709761
H	3.992287488	3.635137152	-2.246458095
H	2.829383850	2.817333111	2.224632402
H	3.773721612	4.247597980	1.715755970
H	-2.184894246	4.587168813	1.845482211
H	-3.576550167	5.122915887	0.867393885
H	-1.930538119	5.448535999	0.309854254
H	-2.707016435	2.265667514	2.678729819
H	-4.304379864	1.517413575	2.516900830
H	-5.080502809	-0.835475679	1.442566201
H	-5.329999045	-0.989339241	-0.302133302
H	-5.340615944	0.600933481	0.444903811
H	-4.485619806	-2.133709472	-1.874607347
H	-4.005471145	-3.831619621	-2.054475042
H	-2.869779641	-2.522748190	-2.487758346
H	-5.169618323	-2.977528572	0.472404428
H	-4.250579637	-4.485545194	0.505801611
H	-1.743800040	-4.793667177	-2.077992330
H	-2.699270800	-5.593263873	-0.809279343
H	2.449237003	-4.387577796	-2.402033505
H	4.857870841	-2.480765675	0.970848370
H	4.145103444	1.897063204	-2.602308012
H	2.022594374	4.375336427	1.980474950
H	-4.160615039	3.288549146	2.474194649
H	-3.891204170	-3.211347984	1.691294159
H	-0.934241116	-5.623239890	-0.735377325
H	-2.989029117	2.706488036	-2.001250961
H	-2.215154064	4.303600252	-1.908302182
H	-3.937213399	4.127594501	-1.479457099

Table G.6. PBE/TZVP-optimized Cartesian coordinates of methylcobalamin.

Co	0.000000000	0.000000000	0.000000000
N	-1.247645839	-1.414464465	0.062159567
N	-1.442299338	1.292981535	0.015697154
N	1.433982319	1.294300888	-0.109560997
N	1.257917671	-1.399702256	0.116311235
C	-0.713265110	-2.808163556	-0.102974231
C	-1.802156884	-3.685769826	0.636571711
C	-3.091947864	-2.809914785	0.461509162
C	-2.526593603	-1.403792082	0.358761610
C	-3.306926731	-0.209115958	0.554748547
C	-2.774250406	1.051370903	0.333159098
C	-3.516635922	2.396192993	0.498190702
C	-2.537616695	3.374046065	-0.220156278
C	-1.238107129	2.621545819	-0.133391126
C	0.000000000	3.252211564	-0.190106966
C	1.241491268	2.635424993	-0.120014787
C	2.552222895	3.393677532	0.029168915
C	3.562090858	2.301746919	-0.426240036
C	2.790182534	1.012471850	-0.164749536
C	3.373183110	-0.241588527	-0.078413342
C	2.574343314	-1.427515372	0.105108981
C	3.104407228	-2.846901740	0.404336818
C	1.805400073	-3.699728518	0.194133004
C	0.691430434	-2.705547539	0.535253436
C	-0.596181543	-3.156641307	-1.594100155
C	-1.975020347	-5.104351670	0.082028724
C	-1.530982908	-3.793075674	2.150480871
C	-4.040324092	-3.155060860	-0.701031954
C	-4.733210612	-0.380415628	1.036323534
C	-4.936895834	2.480752591	-0.078665147
C	-3.530126374	2.759416533	2.002951933
C	-2.848106887	3.703278770	-1.692164214
C	2.761939743	3.686087562	1.534791610
C	2.616397457	4.708030492	-0.754823330
C	3.980765231	2.395812993	-1.903706191
C	4.879896340	-0.356065687	-0.157067007
C	3.592079957	-2.894145640	1.868596124
C	1.780418816	-5.027053219	0.946101690
C	4.197127493	-3.367331786	-0.544686697
N	0.940486946	0.027244341	-4.306997791
C	1.134875767	0.008158800	-2.953777409
N	0.000000000	0.000000000	-2.273966204
C	-2.407025986	-0.017161141	-3.126707995

C	-3.191612527	-0.011575963	-4.282298878
C	-2.583622876	0.016709085	-5.576860660
C	-1.192910998	0.036871024	-5.698855908
C	-0.428621300	0.030765919	-4.530977843
C	-1.010220169	0.008707208	-3.241995256
C	0.032531280	0.169482448	1.975276464
C	-3.439103140	0.017731630	-6.816592526
C	-4.694596809	-0.035507549	-4.162875476
H	-3.681347201	-2.899018284	1.388631634
H	-2.473862493	4.320171012	0.341458062
H	-0.012825734	4.341424315	-0.247421217
H	4.464862093	2.355665268	0.199999872
H	1.754614103	-3.907035530	-0.886569612
H	0.564102559	-2.659223218	1.629478506
H	-1.554360373	-3.027914850	-2.110160943
H	-0.281080846	-4.200246304	-1.725067648
H	0.136416148	-2.507607108	-2.087297359
H	-1.063439582	-5.701273352	0.224514492
H	-2.230177171	-5.124517567	-0.984310187
H	-2.780300930	-5.620295281	0.626066416
H	-1.409904625	-2.808065526	2.625408616
H	-0.637274417	-4.393882034	2.368464190
H	-3.561340744	-3.053867356	-1.683795234
H	-4.913828864	-2.490157880	-0.693282604
H	-5.009795478	0.385331297	1.769051568
H	-4.867528921	-1.346516387	1.534671997
H	-5.471317488	-0.333956126	0.220557645
H	-4.996304712	2.074809303	-1.097907427
H	-5.246483772	3.535632581	-0.119486956
H	-5.674912455	1.960384160	0.540940604
H	-4.099759811	2.031914601	2.596966622
H	-2.509534452	2.801539007	2.411218960
H	-2.918806374	2.793902449	-2.306152075
H	-3.786972389	4.264808346	-1.786851331
H	1.962254470	4.332937987	1.921224754
H	3.723363644	4.200214244	1.687522868
H	2.768748719	2.758782788	2.126437207
H	2.373594285	4.574192659	-1.817225187
H	3.621933746	5.147742757	-0.680515229
H	1.915756722	5.445701847	-0.337301320
H	4.625514238	1.551478973	-2.183243536
H	3.101804927	2.392028255	-2.565799609
H	5.365175173	0.616150779	-0.027549722
H	5.221580867	-0.764793613	-1.120970993

H	5.281710188	-1.011886033	0.625493745
H	3.948489912	-3.902906476	2.124330655
H	2.794011072	-2.624453268	2.576408723
H	4.427357969	-2.200098491	2.034316936
H	1.714772868	-4.885069234	2.034175769
H	0.928099108	-5.645851421	0.636448939
H	4.275176580	-4.460033807	-0.436258507
H	5.189220608	-2.954882208	-0.329816050
H	2.124226161	0.000388392	-2.507034908
H	-2.879729627	-0.047987714	-2.145051875
H	-0.726307967	0.053330777	-6.685983718
H	-0.899734927	-0.243904834	2.380448153
H	0.105023676	1.239150921	2.210765099
H	0.904175671	-0.359480936	2.380976003
H	-4.073361886	-0.881407009	-6.871372351
H	-4.119484617	0.883563782	-6.841689476
H	-2.823044828	0.049729035	-7.724369800
H	-5.007391151	-0.059737983	-3.110420085
H	-5.155671225	0.849117487	-4.631384152
H	-5.131119774	-0.915818944	-4.661529045
H	-2.381218569	-4.291955918	2.638587410
H	-4.414862123	-4.183447657	-0.604326019
H	-3.997660298	3.746297631	2.141441570
H	-2.043122472	4.324022426	-2.109839808
H	4.542789509	3.320382268	-2.094224840
H	2.687461018	-5.613841431	0.738597689
H	3.952986234	-3.157246134	-1.596719622
H	1.670027222	0.045620137	-5.012534548

Table G.7. PBE/TZVP-optimized Cartesian coordinates of adenosylcobalamin.

Co	0.000000000	0.000000000	0.000000000
C	-0.730739382	-2.802331516	-0.162250938
C	-1.828634745	-3.687915758	0.558615738
C	-3.111376253	-2.799103715	0.401713545
C	-2.535013289	-1.395568657	0.326643584
C	-3.308401775	-0.196759930	0.525120504
C	-2.775337766	1.059305607	0.274564954
C	-3.511652913	2.410120780	0.411464465
C	-2.539849923	3.364994550	-0.348100651
C	-1.239148416	2.615886859	-0.244332348
C	0.000000000	3.239528525	-0.318589712
C	1.243063729	2.623945145	-0.224368388

C	2.552759687	3.392832758	-0.133715077
C	2.788499747	0.996347700	-0.205934947
C	3.365988344	-0.258793505	-0.106386328
C	2.564885086	-1.441810707	0.075247160
C	3.092911334	-2.876616609	0.295766647
C	1.785105882	-3.712015406	0.073918285
C	0.678495518	-2.730972903	0.476273160
C	-0.607492968	-3.113930274	-1.661663925
N	-1.254265869	-1.410808951	0.036738700
N	-1.445598219	1.291617060	-0.054710728
N	1.434689980	1.286089255	-0.144648051
N	1.247812829	-1.407998243	0.120907346
C	-2.014806407	-5.090208181	-0.031830496
C	-1.558647075	-3.836929276	2.069870262
C	-4.049572794	-3.113184837	-0.776594066
C	-4.729130849	-0.360422135	1.026381057
C	-4.941898043	2.482196095	-0.141521376
C	-3.504201669	2.817026804	1.904684343
C	-2.869783793	3.643713337	-1.826286121
C	2.610411638	4.641325218	-1.021197445
C	2.757657185	3.807578129	1.341962500
C	3.569231098	2.273225856	-0.499807157
C	4.047741850	2.297981255	-1.962810629
C	4.874821611	-0.371219135	-0.157693031
C	3.625309788	-3.021403582	1.737699750
C	4.155505397	-3.348526361	-0.711925663
C	1.767758971	-5.073586952	0.762634499
N	0.931330100	0.110509451	-4.343512455
C	1.131298185	0.052622804	-2.991667898
N	0.000000000	0.000000000	-2.308504234
C	-2.409433014	-0.053528877	-3.155008275
C	-3.197909580	-0.035963107	-4.307814331
C	-2.596457209	0.053711747	-5.602065226
C	-1.207007365	0.115020632	-5.727118571
C	-0.438276258	0.090863768	-4.562564085
C	-1.014080302	0.016307494	-3.272379934
C	-4.698554271	-0.112994058	-4.183707743
C	-3.457266100	0.075471809	-6.837850254
N	7.430153919	2.725890346	4.137922358
C	7.057302728	1.496694264	3.748069977
N	5.836480477	0.932570875	3.792400117
C	4.936416200	1.779507141	4.309548607
C	5.162044335	3.088663830	4.755961723
C	6.494201026	3.554107946	4.651738701

N	3.993957444	3.672874082	5.220546892
C	3.081735072	2.738618927	5.061064834
N	3.581684802	1.555558125	4.527588521
N	6.863641975	4.798019796	5.030400176
C	2.867911171	0.369592330	4.124388265
C	1.874926295	-0.214213671	5.171974175
C	0.516988745	-0.187735198	4.428882228
C	0.964971882	-0.216986475	2.941275246
C	-0.100755597	0.259884632	1.994774666
O	2.104591730	0.665897462	2.944613864
O	2.161858578	-1.592133404	5.450856483
O	-0.360368205	-1.224013930	4.806774322
H	-3.704308674	-2.903274791	1.324767103
H	-2.475588806	4.328514305	0.182612804
H	-0.009437915	4.325190852	-0.420222109
H	1.713788255	-3.869735407	-1.014635286
H	0.550748507	-2.740744432	1.570849367
H	-0.345241493	-4.168992198	-1.816816202
H	0.174424783	-2.493968173	-2.113684962
H	-1.541534135	-2.910977803	-2.197977439
H	-1.112575229	-5.702045693	0.102822035
H	-2.259766904	-5.080869966	-1.100478849
H	-2.831938093	-5.608487331	0.492127911
H	-0.662075269	-4.441285008	2.264356463
H	-1.445766119	-2.871403710	2.585050985
H	-4.904300546	-2.423818588	-0.784111203
H	-3.543961700	-3.025156735	-1.746152001
H	-4.864557576	-1.331791210	1.515634295
H	-5.478676689	-0.296435424	0.222472652
H	-4.986742567	0.398318703	1.774283513
H	-5.029397758	2.027714593	-1.137760983
H	-5.242579787	3.537204714	-0.225881085
H	-5.669510051	2.001300466	0.520384270
H	-4.074448254	2.113674158	2.525831570
H	-2.478252197	2.863444485	2.299967872
H	-3.821941004	4.181620199	-1.929174156
H	-2.925355061	2.713522145	-2.409918362
H	1.900740331	5.401503959	-0.664268422
H	2.377200756	4.427064870	-2.072533522
H	3.610770233	5.096902651	-0.973591734
H	2.733938791	2.931052240	2.002781544
H	1.972309914	4.505792825	1.666604820
H	3.729501015	4.311765474	1.456374082
H	5.257541350	-1.056259257	0.606560938

H	5.353483160	0.593391059	0.036550279
H	5.243143811	-0.731555916	-1.130965007
H	2.843213969	-2.843198290	2.489789882
H	4.443235353	-2.317645752	1.940301611
H	4.017130491	-4.036402452	1.895561101
H	5.158510391	-2.965078588	-0.493476822
H	4.218441510	-4.447301422	-0.678557414
H	1.737370125	-4.986231006	1.857543785
H	0.901105866	-5.669828784	0.449776761
H	2.122984002	0.049385474	-2.548889345
H	-2.881848487	-0.127287558	-2.174795250
H	-0.744917291	0.176217822	-6.714859497
H	-5.107741441	-0.991095059	-4.708031573
H	-5.190470981	0.770530783	-4.621217929
H	-5.002595240	-0.179891870	-3.130489153
H	-4.070207810	-0.835805828	-6.920837805
H	-2.846643506	0.152327295	-7.746795332
H	-4.157237808	0.926064164	-6.829568410
H	7.857891642	0.871263061	3.342979611
H	2.027775112	2.850822947	5.304087362
H	7.846748146	5.044563588	5.003495492
H	6.199176092	5.408789402	5.491004868
H	3.642183271	-0.386944228	3.907845446
H	1.858861505	0.372206914	6.102555045
H	0.010833143	0.773847498	4.617596886
H	1.292653322	-1.251822144	2.729154953
H	-0.202721111	1.349330572	2.111306084
H	-1.049706580	-0.224923060	2.265027326
H	2.757691453	-1.646543916	6.216039263
H	0.213285283	-1.940509121	5.151171434
H	-4.454321021	-4.130101511	-0.691415002
H	-3.960264424	3.812234993	2.019506545
H	-2.083509207	4.270514798	-2.269895527
H	3.196670085	2.277582712	-2.659736150
H	4.630552183	3.205630090	-2.170212793
H	4.689229860	1.434266348	-2.180751754
H	-2.407433393	-4.359326364	2.535599005
H	1.657954737	0.160126042	-5.050595269
H	4.447868050	2.347867025	0.158332767
H	2.662932046	-5.656473844	0.499749815
H	3.888695432	-3.063832945	-1.740731735

Table G.8. PBE/TZVP-optimized Cartesian coordinates of imidazolylcobalamin.

Co	0.000000000	0.000000000	0.000000000
C	-0.711690806	-2.818756609	-0.145068423
C	-1.811810749	-3.708765067	0.567284582
C	-3.101730499	-2.825083181	0.400868368
C	-2.539930048	-1.415515552	0.319299025
C	-3.325713375	-0.217956426	0.525508148
C	-2.793868315	1.048520344	0.320017679
C	-3.538181021	2.394876030	0.467650540
C	-2.537557760	3.378723272	-0.216638397
C	-1.240885857	2.625492617	-0.118306588
C	-0.000000000	3.252045961	-0.174828642
C	1.245104165	2.640915725	-0.110198843
C	2.555816694	3.403969725	-0.000911807
C	3.568156222	2.304919253	-0.441001449
C	2.804637684	1.014462408	-0.166200329
C	3.385012551	-0.245069899	-0.090125955
C	2.583049614	-1.436470965	0.091020383
C	3.108378603	-2.859188635	0.386842542
C	1.807350640	-3.714104145	0.182878552
C	0.687596015	-2.720733298	0.515856339
C	-0.570667245	-3.148902655	-1.638800700
N	-1.259823453	-1.431268711	0.037798919
N	-1.456203933	1.298905924	0.031748933
N	1.449445304	1.301968513	-0.085622598
N	1.268768286	-1.416813402	0.107519636
C	-1.983020040	-5.111220714	-0.027251186
C	-1.552217514	-3.856352915	2.080044296
C	-4.048491336	-3.153416314	-0.767590346
C	-4.762995953	-0.399153143	0.972231050
C	-4.940795196	2.477804063	-0.150915378
C	-3.598949804	2.757984926	1.972447289
C	-2.818641764	3.731232409	-1.690794657
C	2.783389598	3.744401198	1.492416545
C	2.596274729	4.697262155	-0.822544175
C	3.994216382	2.380234101	-1.918439257
C	4.891012692	-0.365920911	-0.192387218
C	3.597585291	-2.901028195	1.851653124
C	1.778482637	-5.036933813	0.944187394
N	0.944869211	-0.041472616	-4.038394539
C	1.147160767	-0.049702316	-2.699699153
N	-0.000000000	-0.000000000	-2.022923529
C	-2.422912278	0.046480051	-2.910745802
C	-3.181410754	0.061490380	-4.084157000

C	-2.549283466	0.059546283	-5.368825484
C	-1.157830505	0.031576721	-5.454493705
C	-0.421738738	0.013031931	-4.269092110
C	-1.023188340	0.030500254	-2.989302671
C	-4.685354312	0.078333777	-3.994653743
C	-3.375499870	0.084107619	-6.625153805
N	0.083959399	0.098974329	1.992084057
C	1.245756419	0.182579452	2.745709781
C	0.914445589	0.248908120	4.074836357
N	-0.462356246	0.204595809	4.118438492
C	-0.935860510	0.113745234	2.852992560
C	4.198767235	-3.377329632	-0.565562239
H	-3.692737106	-2.921551239	1.325614485
H	-2.482202235	4.316385962	0.359183164
H	-0.013007271	4.339506407	-0.243190673
H	4.467653801	2.367115529	0.188719274
H	1.757355039	-3.929373079	-0.896996630
H	0.542977045	-2.670404171	1.607187479
H	-0.266790750	-4.194961326	-1.771259211
H	0.187131262	-2.513486293	-2.110296691
H	-1.515203479	-3.002557909	-2.172485936
H	-1.069877233	-5.710409378	0.087171567
H	-2.250617569	-5.101297451	-1.090079639
H	-2.781013620	-5.645829658	0.508781222
H	-0.662270018	-4.464563478	2.288332382
H	-2.407287809	-4.366472691	2.546626790
H	-1.433203684	-2.885077862	2.586129412
H	-4.436514522	-4.175795088	-0.670667701
H	-3.564903201	-3.061427631	-1.747920078
H	-4.914956626	-2.479059005	-0.763319509
H	-4.900779991	-1.358559010	1.483964758
H	-5.476101714	-0.378814903	0.134095588
H	-5.074318053	0.378927376	1.677888649
H	-5.693469585	1.946630131	0.440180705
H	-4.968753059	2.090256657	-1.177653637
H	-5.253989256	3.531740207	-0.183794739
H	-4.183323076	2.029381041	2.551558513
H	-4.080317128	3.738861324	2.095580693
H	-2.591347557	2.822861344	2.410341518
H	-2.864699419	2.832239362	-2.321576439
H	-2.020220658	4.376969447	-2.080490111
H	-3.764475721	4.278493382	-1.789562626
H	2.796890982	2.839822108	2.118925813
H	1.995978378	4.411408581	1.869843687

H	3.749165809	4.257485882	1.610669097
H	1.899061228	5.441354797	-0.412069611
H	2.340736645	4.538390985	-1.878283997
H	3.598412149	5.146068738	-0.773928343
H	4.668714539	1.553361520	-2.177040639
H	4.531075381	3.316103171	-2.121289385
H	3.123705221	2.338012485	-2.590074250
H	5.306449808	-1.004054232	0.597030037
H	5.380327235	0.607589728	-0.097018177
H	5.210274533	-0.798862822	-1.152125876
H	3.982872143	-3.901261287	2.091576603
H	2.792760268	-2.667752650	2.564908453
H	4.414378828	-2.187739458	2.026510411
H	2.673676177	-5.636945377	0.726289825
H	0.915754411	-5.647303179	0.648945746
H	1.729581272	-4.891843879	2.032323735
H	1.676135723	-0.068480196	-4.743782182
H	2.134568834	-0.093485403	-2.254886514
H	-2.932010586	0.045980660	-1.949375005
H	-0.661911209	0.020401824	-6.427193503
H	-5.131639184	-0.789945353	-4.504359392
H	-5.109990797	0.973378865	-4.475867498
H	-5.024129359	0.063852464	-2.949984386
H	-4.055455141	-0.780307159	-6.678679859
H	-2.742563795	0.068594651	-7.521663091
H	-4.007029937	0.985383199	-6.675164819
H	2.228477502	0.191177725	2.292717232
H	1.524860314	0.325088896	4.967261232
H	-1.033929895	0.238231833	4.959162376
H	-1.987161414	0.068052028	2.595145103
H	4.267913009	-4.470136147	-0.463601949
H	5.193263670	-2.975531776	-0.344281580
H	3.956106015	-3.160873757	-1.616320994

Table G.9. PBE/TZVP-optimized Cartesian coordinates of aquocobalamin.

Co	0.000000000	-0.000000000	0.000000000
O	0.035411281	0.032388434	2.128472742
C	-0.706483647	-2.820989544	-0.031638594
C	-1.787273847	-3.692459875	0.735911676
C	-3.076407052	-2.806730287	0.603432192
C	-2.509793702	-1.402905127	0.497966056
C	-3.271365596	-0.203499313	0.766812784

C	-2.755867631	1.058900051	0.519282212
C	-3.479429881	2.407098752	0.720593617
C	-2.539441358	3.384615733	-0.055485927
C	-1.238154000	2.635420830	-0.039294313
C	0.000000000	3.263384780	-0.138558048
C	1.242126308	2.651201239	-0.056109003
C	2.553495027	3.412014429	0.062512710
C	3.562182585	2.312764122	-0.382942854
C	2.803445100	1.024607802	-0.092497353
C	3.386627945	-0.229747987	-0.012583257
C	3.113516143	-2.830931444	0.512336795
C	1.818770849	-3.694565897	0.307944537
C	0.695047695	-2.699905839	0.615494252
C	-0.589455399	-3.194011044	-1.516103367
N	-1.251697229	-1.428110043	0.125439924
N	-1.437123672	1.307255460	0.131153504
N	1.442920424	1.307434280	0.001344353
N	1.272473662	-1.400041549	0.180016863
C	-1.983865885	-5.103704764	0.171440630
C	-1.486093831	-3.811005583	2.243695991
C	-4.047860941	-3.120273555	-0.549488896
C	-4.673254849	-0.382319719	1.312490028
C	-4.930229255	2.491993498	0.226619314
C	-3.405759541	2.779486094	2.222863716
C	-2.932941344	3.715061555	-1.508623174
C	2.600560386	4.711260615	-0.748546647
C	3.966598781	2.378501505	-1.866240635
C	4.891787764	-0.348347289	-0.109793821
C	3.590267631	-2.839382646	1.981880885
C	4.214716954	-3.360678148	-0.420736506
C	1.790822302	-5.005835014	1.086572413
C	1.136447335	-0.047564032	-2.632799642
N	0.000000000	0.000000000	-1.935224929
C	-2.445852184	0.042342853	-2.756146613
C	-3.231643339	0.061080832	-3.910982951
C	-2.628288227	0.070806059	-5.210188596
C	-1.239057805	0.048945275	-5.329760087
C	-0.473912266	0.025044993	-4.162919590
C	-1.049821622	0.034682436	-2.872093895
C	-4.732429509	0.069221363	-3.786363501
C	-3.483945571	0.096678633	-6.446097760
C	2.586633680	-1.418193611	0.182052727
N	0.898596290	-0.029207901	-3.963516106
C	2.779058683	3.736885861	1.560671922

H	-3.644828815	-2.908105675	1.540006054
H	-2.444860245	4.330473441	0.499878043
H	-0.011451942	4.349633759	-0.222617866
H	4.469446399	2.377200433	0.235090182
H	1.777466585	-3.924876333	-0.769832951
H	0.554404962	-2.618243544	1.704876374
H	-1.542068046	-3.060785623	-2.039616898
H	-0.292150940	-4.244632567	-1.622852825
H	0.165083281	-2.579410881	-2.021312437
H	-1.073315497	-5.708782192	0.274479680
H	-2.276593298	-5.109897734	-0.885098502
H	-2.772629212	-5.622285621	0.736092447
H	-0.579896128	-4.397464370	2.441685964
H	-1.380139031	-2.831200703	2.735898724
H	-3.585267146	-3.003160668	-1.537771762
H	-4.420954896	-4.149119770	-0.463143280
H	-4.932183636	0.391877513	2.042657040
H	-4.780069946	-1.341674563	1.829391159
H	-5.440063022	-0.354942861	0.523075776
H	-5.053648266	2.075951749	-0.782109278
H	-5.235189414	3.547952799	0.192981566
H	-5.630869023	1.985983010	0.897159194
H	-2.363598927	2.847659155	2.571519340
H	-3.871869319	3.762849738	2.380036805
H	-3.017580329	2.809720628	-2.125888925
H	-3.888378089	4.254067472	-1.540835520
H	2.330753550	4.562694267	-1.802527534
H	3.609854248	5.143833660	-0.709734414
H	1.919877200	5.462977278	-0.325186110
H	3.085032332	2.346776333	-2.524713384
H	4.625200387	1.541016512	-2.132197810
H	5.212968453	-0.805783558	-1.057568554
H	5.302669785	-0.967907729	0.696718331
H	5.379809628	0.627907974	-0.036127423
H	4.423175325	-2.140678160	2.140154420
H	3.951845061	-3.840968262	2.251300698
H	2.788556524	-2.566732380	2.684103997
H	5.207992699	-2.960179738	-0.191845015
H	4.279715141	-4.452440390	-0.307077170
H	1.733438144	-4.846078934	2.172200309
H	2.693547289	-5.600214903	0.884377072
H	2.131386121	-0.094986899	-2.204821430
H	-2.927620049	0.027218561	-1.781190417
H	-0.767162038	0.045539360	-6.314422285

H	-5.186085427	-0.789416646	-4.305259530
H	-5.169988601	0.973071001	-4.238471027
H	-5.048087533	0.031791409	-2.734899299
H	-2.872381605	0.121435643	-7.356943700
H	-4.144744180	0.977308998	-6.460476651
H	-4.136651577	-0.788476593	-6.501889447
H	-2.319092555	-4.332177470	2.737080623
H	-4.922327771	-2.456229819	-0.511201112
H	-3.934476762	2.055241368	2.856589101
H	-2.173288626	4.364332385	-1.964622040
H	4.511970082	3.306389522	-2.081523564
H	3.987421248	-3.153938688	-1.476976309
H	0.934552583	-5.625384503	0.792250739
H	1.612952444	-0.057190106	-4.686479206
H	2.818129623	2.823446645	2.174682124
H	1.982380079	4.385343432	1.950941788
H	3.735733159	4.264694200	1.685676934
H	-0.839645649	0.312417293	2.462932042
H	0.660778341	0.726965712	2.412724690

Table G.10. PBE/TZVP-optimized Cartesian coordinates of hydroxocobalamin.

Co	0.000000000	0.000000000	-0.000000000
O	0.109366254	0.009011919	1.899442558
C	-0.689347749	-2.814199304	-0.003103298
C	-1.766869074	-3.668651799	0.777750203
C	-3.066260537	-2.817719878	0.568801514
C	-2.513599389	-1.409505650	0.427993295
C	-3.299094736	-0.215748674	0.630239972
C	-2.771033473	1.048545971	0.420869462
C	-3.507345895	2.392538856	0.621938826
C	-2.536913551	3.383385502	-0.090949996
C	-1.234926338	2.631200067	-0.029336328
C	-0.000000000	3.264523482	-0.066972715
C	1.241740112	2.643424113	0.005833558
C	2.549666509	3.386486583	0.235848475
C	3.559697293	2.338117365	-0.309918715
C	2.797148465	1.033184612	-0.109960429
C	3.385276707	-0.217696089	-0.055528580
C	3.120549538	-2.792261848	0.576224936
C	1.831902704	-3.669937498	0.404747892
C	0.706677861	-2.658865633	0.646130360
C	-0.552099878	-3.226131651	-1.477218782

N	-1.249359257	-1.426283588	0.087759676
N	-1.443840106	1.298992879	0.090308418
N	1.440398215	1.310395622	-0.052840453
N	1.279131416	-1.394251919	0.133190986
C	-1.921027833	-5.111952425	0.289039944
C	-1.497276309	-3.685440119	2.296599019
C	-3.998486859	-3.192116425	-0.598080432
C	-4.723025816	-0.392101540	1.118055793
C	-4.934670426	2.496269194	0.067556269
C	-3.495970563	2.724038380	2.133855505
C	-2.863097289	3.731508784	-1.554698937
C	2.627544873	4.769293697	-0.414952406
C	3.935635311	2.531301349	-1.788758490
C	4.889943339	-0.334520771	-0.163939820
C	3.562613857	-2.705597520	2.053392996
C	4.243143651	-3.381336502	-0.292690962
C	1.806391804	-4.934677956	1.258808754
C	1.131935074	-0.001683273	-2.770270970
N	-0.000000000	0.000000000	-2.076721761
C	-2.425522044	-0.025074321	-2.902467156
C	-3.217555452	-0.029518685	-4.052775974
C	-2.620834148	-0.009244283	-5.352224756
C	-1.231423416	0.007964350	-5.483075021
C	-0.459048825	0.010207126	-4.320948170
C	-1.030231990	0.000241615	-3.027823014
C	-4.718968211	-0.059011858	-3.920970243
C	-3.485939757	-0.013608961	-6.584272454
C	2.589499888	-1.404477860	0.156988509
N	0.912465226	0.005813690	-4.114471207
C	2.736754018	3.524666571	1.766556004
H	-3.660481980	-2.894865996	1.492987977
H	-2.470223397	4.320973643	0.483537456
H	-0.014151557	4.354875292	-0.091343173
H	4.475453755	2.356678096	0.298381750
H	1.806072756	-3.963431127	-0.659204679
H	0.559224121	-2.496304852	1.725044065
H	-1.506631010	-3.125841118	-2.007770701
H	-0.226642391	-4.271357593	-1.560637860
H	0.185605863	-2.593519211	-1.986465556
H	-1.002885155	-5.691121089	0.460769195
H	-2.170318671	-5.183151833	-0.776473281
H	-2.722338319	-5.612742971	0.853166153
H	-0.605190591	-4.271507916	2.553816853
H	-1.373113365	-2.671134972	2.705067614

H	-3.522845285	-3.058045030	-1.580215810
H	-4.326511401	-4.236853858	-0.513487859
H	-5.000855175	0.373197002	1.849942161
H	-4.854241471	-1.357732025	1.619082325
H	-5.463260640	-0.350170922	0.303478150
H	-5.009548119	2.126290519	-0.964184056
H	-5.243229605	3.552467304	0.068348694
H	-5.664680134	1.954880681	0.677560498
H	-2.467336566	2.767376515	2.522512251
H	-3.965007721	3.705207288	2.303329245
H	-2.942829382	2.828756707	-2.177720268
H	-3.803471839	4.293506958	-1.630851957
H	2.385572290	4.747370407	-1.485948403
H	3.637941011	5.189342592	-0.298062669
H	1.936797511	5.469987646	0.075708347
H	3.038751675	2.587493816	-2.424433182
H	4.556525193	1.698615175	-2.147622929
H	5.202751020	-0.880770285	-1.067384629
H	5.323003062	-0.865563619	0.695175099
H	5.372984591	0.646516706	-0.198267186
H	4.406151227	-2.013251264	2.173801985
H	3.892850164	-3.691089627	2.411545592
H	2.750010479	-2.352419631	2.704773992
H	5.230695023	-2.968791526	-0.059074263
H	4.299499583	-4.466582477	-0.114479285
H	1.721191761	-4.709610201	2.331027178
H	2.724895181	-5.522608554	1.111373879
H	2.126910362	-0.008207046	-2.335040009
H	-2.893956862	-0.046272323	-1.919625342
H	-0.768788542	0.016913040	-6.472265142
H	-5.155119399	-0.941604766	-4.415654381
H	-5.187862913	0.821825383	-4.388829488
H	-5.023818690	-0.081227459	-2.866655735
H	-2.878280552	0.008890659	-7.497519169
H	-4.162451630	0.855681968	-6.608592295
H	-4.125787907	-0.909614755	-6.627668663
H	-2.350710810	-4.153744194	2.809163622
H	-4.901022192	-2.566563101	-0.581348576
H	-4.049274066	1.981549430	2.724538128
H	-2.063772565	4.358807050	-1.974898244
H	4.507159842	3.458389546	-1.932067892
H	4.046350152	-3.233650112	-1.365779939
H	0.966758087	-5.586878513	0.983193227
H	1.633088491	0.010658369	-4.829317791

H	2.698875375	2.545254803	2.264177159
H	1.949813589	4.157256872	2.200421068
H	3.709761256	3.991619160	1.982830330
H	-0.675242831	0.499307407	2.205415697

Table G.11. PBE/TZVP-optimized Cartesian coordinates of cob(I)alamin.

Co	0.000000000	0.000000000	0.000000000
C	2.708770974	-0.530691155	-0.640152141
C	3.699756052	-1.749627094	-0.506517246
C	2.806412624	-2.944756088	-0.989157092
C	1.403470032	-2.460268892	-0.664297141
C	0.249475685	-3.297324111	-0.598850825
C	-1.018278429	-2.749876782	-0.474301144
C	-2.342739864	-3.539407428	-0.403072202
C	-3.366615512	-2.423746617	-0.756059996
C	-2.625807683	-1.189524921	-0.314587714
C	-3.264771503	-0.000000000	0.000000000
C	-2.636016096	1.188464836	0.333281562
C	-3.371263176	2.433316324	0.806588039
C	-2.291846361	3.512934538	0.531001631
C	-1.003630223	2.702333320	0.594902514
C	0.248383964	3.257945298	0.803911183
C	2.866753353	2.896052837	1.058456738
C	3.690409248	1.731015538	0.414171631
C	2.718338369	0.548307657	0.466089300
C	2.750865773	0.126824616	-2.030671121
N	1.360966796	-1.151592318	-0.452218556
N	-1.277214864	-1.387175962	-0.302549979
N	-1.281824440	1.361720487	0.341274716
N	1.370033252	1.166337644	0.408369874
C	4.996104312	-1.604954046	-1.312142508
C	4.069732977	-2.038095388	0.962451042
C	2.955065938	-3.397770255	-2.453247868
C	0.466138228	-4.797389593	-0.636445150
C	-2.495710170	-4.759565114	-1.323046879
C	-2.580141019	-3.963741853	1.065590392
C	-3.735642519	-2.304772480	-2.246585376
C	-4.718487533	2.675834243	0.119864631
C	-2.409749355	4.208624758	-0.835908713
C	0.357187294	4.739326931	1.101385891
C	3.035382216	2.945932500	2.592951370

C	3.331066721	4.222590853	0.433137219
C	5.066295503	1.497210689	1.033917347
C	1.428692092	2.455861028	0.726365210
C	-3.592133015	2.320278053	2.333689846
H	3.059656758	-3.817112888	-0.363711367
H	-4.292753225	-2.562756371	-0.173704463
H	-4.355834994	-0.009675599	-0.009717409
H	-2.334476955	4.277641288	1.321292388
H	3.827611280	2.011217903	-0.643643921
H	2.802724816	0.036426458	1.439141226
H	2.543048697	-0.611796454	-2.813773610
H	3.729223749	0.581630360	-2.240849058
H	1.975381976	0.902994790	-2.089057960
H	5.588798222	-0.746802610	-0.964417553
H	4.816877201	-1.472706619	-2.386540381
H	5.622810878	-2.501949935	-1.187639635
H	4.670304790	-1.232707249	1.405570613
H	3.175809255	-2.188772659	1.586619581
H	2.696689309	-2.604110772	-3.167921833
H	3.981505601	-3.732075581	-2.665563018
H	-0.158684374	-5.319045454	0.100335166
H	1.504902805	-5.045671256	-0.389476905
H	0.254681990	-5.245339275	-1.620442517
H	-2.148149052	-4.554832693	-2.345080527
H	-3.559239725	-5.040671872	-1.378226940
H	-1.954512274	-5.634965469	-0.948198061
H	-2.566528157	-3.090549233	1.733772143
H	-3.559994714	-4.457613657	1.163394355
H	-2.836841728	-2.162173684	-2.864696782
H	-4.274139529	-3.191033086	-2.612165432
H	-4.633855086	2.666445294	-0.975360983
H	-5.139984405	3.645280496	0.427196040
H	-5.447214368	1.903223756	0.407020431
H	-2.379372847	3.473405179	-1.653834673
H	-1.574075339	4.906477496	-0.984646324
H	0.729738368	5.315522554	0.239196584
H	1.039928019	4.939916998	1.938023929
H	-0.612303827	5.166422625	1.379664341
H	2.353122944	3.685265893	3.037636769
H	4.061078065	3.233563651	2.868999313
H	2.811043599	1.973877028	3.056316185
H	2.938313280	5.103760617	0.953727379
H	4.429829199	4.283991116	0.481149989
H	4.993736784	1.104482041	2.058803962

H	5.646558824	2.432229114	1.074352294
H	4.669613137	-2.959852178	1.011511730
H	2.284007102	-4.244538741	-2.654550672
H	-1.808116798	-4.666717270	1.410769244
H	-4.383987603	-1.430481109	-2.398307010
H	-3.346029858	4.780049340	-0.916896375
H	3.036267386	4.285770084	-0.624561427
H	5.655766161	0.780805116	0.446106155
H	-2.635801040	2.182442280	2.860162119
H	-4.237352970	1.462312954	2.572288730
H	-4.072063541	3.233334369	2.722525428

Table G.12. PBE/TZVP-optimized Cartesian coordinates of bisimidazolylcobalamin.

Co	0.000000000	0.000000000	0.000000000
C	-0.687989293	-2.826355113	-0.152715451
C	-1.823120279	-3.706688862	0.510847467
C	-3.105206384	-2.844049748	0.219788432
C	-2.551577260	-1.429579037	0.155847988
C	-3.360075638	-0.236586567	0.277973827
C	-2.817577051	1.034001001	0.134346241
C	-3.573334197	2.376872422	0.258314280
C	-2.532731773	3.373886989	-0.340443695
C	-1.240759875	2.623850973	-0.165939251
C	0.000000000	3.252614721	-0.168425973
C	1.244648940	2.639846952	-0.096027617
C	2.559512949	3.398889213	-0.000623084
C	3.559937107	2.296981733	-0.463322769
C	2.799427921	1.009327710	-0.169772136
C	3.380948410	-0.246803744	-0.074884905
C	2.582457563	-1.433942871	0.141706833
C	3.108981328	-2.843846041	0.489218007
C	1.816138227	-3.711022616	0.285298698
C	0.681913759	-2.716877298	0.563106050
C	-0.476976348	-3.171812266	-1.635922825
N	-1.250719297	-1.440182880	-0.010057935
N	-1.460984565	1.295441029	-0.037897773
N	1.446113733	1.301208664	-0.069755734
N	1.268872417	-1.415963916	0.153487499
C	-1.938526039	-5.132099713	-0.038191908
C	-1.668967814	-3.789736084	2.042143509
C	-3.954874175	-3.211745636	-1.010298645
C	-4.831696586	-0.420942373	0.592090089

C	-4.936772669	2.478359630	-0.440828407
C	-3.729106363	2.706003543	1.764192832
C	-2.716644720	3.750237097	-1.822939245
C	2.812431914	3.733723547	1.490106831
C	2.596025273	4.695287060	-0.817291794
C	3.944992476	2.370642310	-1.951637116
C	4.886070800	-0.370313491	-0.181068514
C	3.568269646	-2.836485160	1.964195593
C	1.779161587	-5.007003699	1.090744071
N	0.659491883	-0.067700083	-4.077257788
C	1.063707675	-0.090367432	-2.784235192
N	0.000000000	0.000000000	-1.984128175
C	-0.714101868	0.040268407	-4.110596119
C	-1.115887836	0.079426404	-2.799453752
N	0.006388202	0.112637178	1.983533532
C	1.131321204	0.237868044	2.783830973
C	0.741061428	0.299300642	4.097248763
N	-0.634032842	0.208467965	4.082811476
C	-1.050773636	0.096171421	2.798819513
C	4.221427722	-3.393045630	-0.418026975
H	-3.765664797	-2.929018677	1.098672159
H	-2.519764849	4.304344982	0.248743855
H	-0.011828911	4.340974899	-0.235227912
H	4.475836372	2.353736384	0.142335035
H	1.793625336	-3.964351086	-0.787491026
H	0.499481950	-2.659480086	1.649189730
H	-0.199834793	-4.227802463	-1.744980316
H	0.325431517	-2.566447142	-2.073015206
H	-1.388015088	-3.002910485	-2.221068335
H	-1.030716351	-5.715412950	0.168609561
H	-2.126504949	-5.167798946	-1.117244461
H	-2.769392456	-5.654516946	0.458474554
H	-0.791471943	-4.378937289	2.341057762
H	-2.550529567	-4.290904501	2.467435036
H	-1.596839787	-2.797743201	2.514243391
H	-4.328036048	-4.240387522	-0.922312892
H	-3.398255220	-3.132982588	-1.953264523
H	-4.831477847	-2.556269333	-1.090028480
H	-5.026175064	-1.404968341	1.032529636
H	-5.469759143	-0.336020642	-0.299795179
H	-5.189333211	0.319376446	1.315994358
H	-5.727610034	1.951767674	0.102340948
H	-4.911674204	2.095479179	-1.469409505
H	-5.238407637	3.535234414	-0.483714846

H	-4.351105972	1.966346123	2.287066876
H	-4.217874757	3.684863397	1.875926889
H	-2.751884289	2.758512214	2.267961303
H	-2.734590511	2.863685747	-2.473197702
H	-1.888805139	4.392976707	-2.153288159
H	-3.649408823	4.308051353	-1.974088684
H	2.849669619	2.827606019	2.112412758
H	2.027091577	4.392024684	1.886948757
H	3.774833842	4.256872983	1.589401146
H	1.910817610	5.443086504	-0.393778426
H	2.324600156	4.543245073	-1.870369301
H	3.602524103	5.136569089	-0.780898446
H	4.569668749	1.515083859	-2.241086317
H	4.518772137	3.283147404	-2.159636538
H	3.054549593	2.381405911	-2.598016894
H	5.310370570	-0.927261263	0.664945479
H	5.375118278	0.608057370	-0.185349512
H	5.200564380	-0.893307965	-1.096489611
H	3.925644040	-3.835323120	2.251241040
H	2.756617711	-2.556476626	2.651846768
H	4.399486135	-2.137030818	2.120978311
H	2.695433537	-5.593661597	0.929006647
H	0.942003702	-5.647152148	0.781642153
H	1.685080167	-4.826004864	2.170811706
H	1.272847762	-0.116324433	-4.887266520
H	2.097086656	-0.169744631	-2.466477006
H	2.131198791	0.279327524	2.369728451
H	1.309277143	0.399996998	5.014864180
H	-1.240658327	0.227445774	4.899081928
H	-2.088922163	0.016656578	2.497064924
H	4.262819230	-4.486860756	-0.301422892
H	5.214963076	-3.009470012	-0.163681334
H	4.024751899	-3.184643993	-1.480551034
H	-1.273029894	0.078830598	-5.038623937
H	-2.118449484	0.153728257	-2.397022809

SECTION G.5. CHAPTER G REFERENCES

1. Medek, A.; Frydman, V.; Frydman, L., *Proc Natl Acad Sci USA* **1997,** *94*, 14237-14242.
2. Harris, R. K.; Becker, E. D.; De Menezes, S. M. C.; Goodfellow, R.; Granger, P., *Pure Appl Chem* **2001,** *73*, 1795-1818.
3. Harris, R. K.; Becker, E. D.; De Menezes, S. M. C.; Granger, P.; Hoffman, R. E.; Zilm, K. W., *Pure Appl Chem* **2008,** *80*, 59-84.
4. Cavanagh, J.; Fairbrother, W. J.; Palmer III, A. G.; Skelton, N. J.; Rance, M., *Protein NMR Spectroscopy: Principles and Practice*. Second ed.; Elselvier, Inc. : Burlington, MA, 2007.
5. Kaczka, E. A.; Wolf, D. E.; Kuehl Jr., F. A.; Folkers, K., *J Am Chem Soc* **1951,** *73*, 3569-3573.
6. Firth, R. A.; Hill, H. A. O.; Pratt, J. M., *Biochem* **1967,** *6*, 2178-2189.
7. Hannibal, L.; Bunge, S. D.; van Eldik, R.; Jacobsen, D. W.; Kratky, C.; Gruber, K.; Brasch, N. E., *Inorg Chem* **2007,** *46*, 3613-3618.
8. Marques, H. M.; Marsh, J. H.; Mellor, J. R.; Munro, O. Q., *Inorga Chim Acta* **1990,** *170*, 259-269.
9. Hill, J. A.; Pratt, J. M.; Williams, R. J. P., *J Chem Soc* **1964**, 5149-5153.
10. Toohey, J. I., *Biochem* **1965,** *54*, 934-942.
11. Liptak, M. D.; Brunold, T. C., *J Am Chem Soc* **2006,** *128*, 9144-9156.
12. Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Lidwig, M. L., *Science* **1994,** *266*, 1669-1674.
13. Fabbiani, F. P. A.; Buth, G.; Dittrich, B.; Sowa, H., *Crystengcomm* **2010,** *12*, 2541-2550.